Modeling and Pricing Cyber Insurance
–
A Survey

Kerstin Awiszus[1,4], Thomas Knispel[2], Irina Penner[3], Gregor Svindland[1], Alexander Voß[1], and Stefan Weber[1]

[1]House of Insurance, Leibniz Universität Hannover
[2]Berlin School of Economics and Law
[3]HTW, University of Applied Sciences, Berlin
[4]Hannover Rück SE, Hannover

December 17, 2021

The paper presents a comprehensive overview of modeling and pricing cyber insurance including clear and easily understandable explanations of the underlying mathematical concepts. We distinguish between three main types of cyber risks: idiosyncratic, systematic, and systemic cyber risks. While for idiosyncratic and systematic cyber risks, classical actuarial modeling approaches seem well-suited, systemic cyber risks require a more sophisticated approach, e.g., based on epidemic network models. In the context of pricing cyber insurance, interdependence issues arise for both systematic and systemic cyber risks. In this case, classical actuarial valuation has to be replaced by a more complex analysis, e.g., based on the concepts of risk-neutral valuation and (set-valued) monetary risk measures.

Keywords: Cyber Risks; Cyber Insurance; Idiosyncratic Risk; Systematic Risk; Systemic Risk.

1. Introduction

Cyber risks constitute a major threat to companies worldwide.[1] In the last years, the estimated costs of cyber crime have continuously been increasing – from approximately USD 600 billion in 2018 to more than USD 1 trillion in 2020, cf. CSIS (2020). Consequently, the market for cyber insurance is experiencing a strong growth providing contracts that mitigate the increasing risk exposure. However, cyber insurance differs from other lines of business in multiple ways posing significant challenges to insurance companies that offer cyber coverage:

- *Data* of cyber events and losses are scarce and typically not available in the desired amount or granularity.
- Cyber threats are evolving dynamically in a highly *non-stationary* cyber risk landscape.

[1]For example, according to the annually published Allianz Risk Barometer (see, e.g., Allianz (2021)), cyber risk ranges among the top three global business risks since 2016.

- *Aggregate cyber risks* arise due to common IT architectures or complex interconnections that cannot easily be captured.
- The term 'cyber' risk itself comprises many *different types of risks* with different root causes and types of impact.

Insurance companies cannot solely rely on standard actuarial approaches when modeling and pricing cyber risks. These methods need to be complemented by novel and innovative techniques for both underwriting and quantitative risk management. The current paper provides the following main contributions:

(i) We provide a *comprehensive overview of the state of the art of modeling and pricing cyber insurance*. In contrast to other surveys (see, e.g., Eling (2020)) that focus on a high-level review of the literature, we present and explain the underlying mathematical concepts and discuss their advantages and drawbacks.[2]

(ii) The second main contribution of the paper is a novel classification of cyber risks into three different types: *idiosyncratic, systematic, and systemic cyber risks*. While the distinction between idiosyncratic and systemic risks is common in the current cyber insurance literature (see, e.g., Zeller and Scherer (2021)), a further refinement is necessary. Systemic risk refers to risk that arises from the interaction of the components of a system, either on a local or on a global level.

Idiosyncratic and systematic cyber risks can be captured by classical actuarial approaches; systemic cyber risks require different methodologies such as epidemic network models which capture the interconnectedness of the entities. We suggest pricing techniques that adequately incorporate interdependence for both systematic and systemic cyber risks by combining the concepts of risk-neutral valuation and risk measures.

The paper is structured as follows. Section 2 reviews classical actuarial approaches. We begin with an introduction to the frequency-severity approach in the context of cyber risk and discuss how to model both idiosyncratic and systematic risks in this framework. We explain how to capture dependence in such models. Systemic cyber risks are considered in Section 3. In the first part, we explain how network models capture the interconnections of entities. In the second part, we discuss game-theoretic approaches that focus on the strategic interaction of agents. In Section 4, we describe pricing methods for cyber insurance contracts that are applicable in the face of idiosyncratic, systematic, and systemic risks. Section 5 discusses open questions for future research.

2. Classical Actuarial Approaches Applied to Cyber Risks

The pricing of cyber insurance contracts as well as quantitative cyber risk management require a sound model for the loss distribution, customized to the application purpose. While classical actuarial premium principles are essentially related to the expected claims amount (plus a safety loading), quantitative risk management particularly refers to extreme losses in the tail of the distribution and their quantification in terms of risk measures such as *Value at Risk* or *Average Value at Risk*, see Section 4.

In actuarial mathematics, a standard model for insurance losses – used across all lines of business – is the *frequency-severity approach*, also called *collective risk model*. For a certain time interval $[0, t]$, $t > 0$ (typically $t = 1$ year), a collective of policyholders causes a random

[2]Surveys that include detailed conceptual explanations are, e.g., Böhme and Schwartz (2010), Marotta et al. (2017), and Böhme, Laube, and Riek (2018). In contrast to our paper, these authors focus exclusively on a game-theoretic models. We discuss such a perspective in Section 3.2.

number of claims $\mathcal{N}_t$ (*frequency*) with corresponding random loss sizes $\mathcal{Y}_1, \mathcal{Y}_2, \ldots$ (*severity*) generating the total claims amount

$$\mathcal{S}_t = \sum_{j=1}^{\mathcal{N}_t} \mathcal{Y}_j, \quad t > 0.$$

Calculations within the frequency-severity approach typically rely on the following mathematical assumptions (see, e.g., Mikosch (2004)):

(C1) Claims occur at arrival times $0 \leq T_1 \leq T_2 \leq \ldots$. The number of claims in the time interval $[0, t]$, $t \geq 0$, is defined by

$$\mathcal{N}_t := \#\{j \geq 1 \mid T_j \leq t\},$$

i.e., $\mathcal{N} = (\mathcal{N}_t)_{t \geq 0}$ constitutes a counting process on $[0, \infty)$.

(C2) The jth claim arriving at time T_j causes the claim size $\mathcal{Y}_j$. It is assumed that the sequence $(\mathcal{Y}_j)_{j \geq 1}$ of claim sizes consists of independent and identically distributed random variables.

(C3) Claim sizes and claim numbers are assumed to be independent from each other.

In contrast to classical insurance risks, however, cyber risk is more challenging in different ways. In particular, the standard assumptions of the frequency-severity approach as well as classical statistical techniques are no longer applicable:

- Claims *data* are not available in sufficient quantity or in the required granularity.
- Technology and cyber threats are evolving rapidly, i.e., the cyber environment is highly *non-stationary*.
- Cyber incidents may affect different policyholders at the same time, i. e., the typical assumption of *independence* for insurance risks does not hold any longer. Moreover, there is – in contrast to natural catastrophe risks – no simple geographical delimitation of dependent risks.

Nonetheless, the frequency-severity approach can be customized to account for cyber risk – at least in first proximity and for certain types of **non-systemic** cyber risks, which can be subdivided into

- **Idiosyncratic risks**: cyber risks that occur at individual policyholders – independently of the other firms; thus, they are subject to pooling of risk. For example, targeted/tailor made attacks (including individual failure) on individual firms. Modeling of such incidents highly depends on the firm's individual characteristics.
- **Systematic risks**: cyber risks resulting from common vulnerabilities of the insured; therefore, they affect different firms at the same time, e.g., due to utilization of the same software, server, or computer system. These risks can be modeled via common risk factors.

In the frequency-severity approaches presented below, we thus explicitly distinguish between techniques suitable for modeling idiosyncratic or systematic incidents. In the context of cyber insurance, however, a third class of risks can be identified, namely

- **Systemic risks**: cyber risks resulting from being a part of a network; for example, malware or supplier attacks.

Proper modeling of such risks goes beyond the classical framework of actuarial modeling and requires appropriate models for networks, disease spread, and strategic interaction. Hence, we discuss the modeling of systemic cyber risks separately in Section 3, while the pricing for all types of cyber risks is discussed in Section 4.

To adopt the frequency-severity approach in the context of cyber risk, we consider an insurer's portfolio of n policyholders (firms) exposed to the considered type of cyber risk incidents. Each firm admits an individual risk profile characterized by a vector of covariates, e.g., *industry sector*, *size*, *IT security level*, which are elicitable, for example, via a questionnaire or from public information. Using the covariates, the insurer's portfolio is decomposed into homogeneous groups, labeled $\{1, \dots, K\}$, with covariates vector x^k for group k. We denote by n_k, $k = 1, \dots, K$, the number of firms in group k, i.e., $n_1 + \ldots + n_K = n$. For pricing purposes, these homogeneous groups can be viewed as tariff cells, i.e., the insurance firm should charge all firms within group k the same premium π_k. In particular, if n_k is large, then the premium of the idiosyncratic cyber risk can be derived from the law of large numbers as the expected claims amount per firm of group k plus a suitable safety loading to avoid ruin in the long run.

Both idiosyncratic and systematic incidents can be grouped into different cyber risk categories, labeled $\{1, \dots, C\}$. Categories may include, for example, *data breach*, *fraud*, and *business interruption*. Two exemplary actuarial classification approaches are sketched and discussed in Appendix A. Cyber risk is modeled per risk category $c \in \{1, \dots, C\}$ and per group $k \in \{1, \dots, K\}$. A pair $m := (c, k)$ is called a *cyber risk module*. The total number of modules $C \cdot K$ is a trade-off between homogeneity and availability of data for statistical estimation.

Within this framework, we model the losses for an insurance company – for each and every cyber risk module as well as on an aggregate level. For this purpose, we first focus on frequency-severity based approaches to modeling cyber risks in the spirit of the classical collective risk model. Second, we add dependence to our cyber risk model in order to capture accumulation risks. Note that appropriate dependence modeling is particularly important for calculating capital requirements in quantitative risk management, since the underlying risk measures refer to events in the extreme tail of the loss distribution.

2.1. Frequency and Severity

A frequency-severity model may be applied on the level of each cyber risk module $m = (c, k)$. For simplicity, we describe the losses per risk category of individual firms by a collective risk model. This can be justified as follows: Since all firms in any group are homogeneous, they will be charged the same premium for any given risk category. From the point of view of the insurance company, only aggregate losses are relevant, i.e., an artificial allocation of losses to individual companies for pricing purposes will produce the correct implications. We thus describe the losses per risk category at the level of any individual firm by a collective risk model with the same severity as the corresponding module, but with a suitably reduced frequency.

For a firm i in group k and a fixed risk category c, i.e., a cyber risk module $m = (c, k)$, we consider the frequency and severity model $(\mathcal{N}^{m,i}, (\mathcal{Y}_j^{m,i})_{j \geq 1})$. Then the total claims amount of firm i up to time t can easily be obtained by summing up:

$$\mathcal{S}_t^{m,i} = \sum_{j=1}^{\mathcal{N}_t^{m,i}} \mathcal{Y}_j^{m,i}.$$

In mathematical terms, all quantities correspond to random variables on a suitable probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbb{P}$ plays the role of the statistical measure.

As outlined in the introduction of this section, one of the most common assumptions in the frequency-severity model is assumption (C3), i. e., claim numbers and sizes are independent of each other. This assumptions facilitates and simplifies many calculations regarding the compound total claim amount process. In particular, the expected total claims amount and its

variance follow from *Wald's formulas:*

$$\mathbb{E}[\mathcal{S}_t^{m,i}] = \mathbb{E}[\mathcal{N}_t^{m,i}] \cdot \mathbb{E}[\mathcal{Y}_1^{m,i}], \quad \operatorname{Var}(\mathcal{S}_t^{m,i}) = \mathbb{E}[\mathcal{N}_t^{m,i}] \operatorname{Var}(\mathcal{Y}_1^{m,i}) + \operatorname{Var}(\mathcal{N}_t^{m,i})(\mathbb{E}[\mathcal{Y}_1^{m,i}])^2.$$

However, the independence assumption may not always be reasonable. For example, Sun, Xu, and Zhao (2020) detect a positive nonlinear dependence between frequency and severity in hacking breach risks on firm-level. More precisely, a firm with a strong cyber self protection is expected to experience both less and weaker hacking attacks than companies with weak self protection mechanisms. In mathematical terms, the authors capture this dependence between frequency and severity by the Gumbel copula, see also Section 2.2.2.

2.1.1. Frequency

Let $\mathcal{N}_t^{m,i}$ denote the number of incidents in module $m = (c, k)$ until time t that are allocated to a firm i in group k, and let $(\mathcal{N}_t^{m,i})_{t\geq 0}$ denote the corresponding counting process. At the aggregate level,

$$\mathcal{N}_t^{m,agg} := \sum_{i=1}^{n_k} \mathcal{N}_t^{m,i} \text{ and } \mathcal{N}_t^{(c)} := \sum_{k=1}^{K} \mathcal{N}_t^{m,agg}, \quad t \geq 0,$$

will count the total number of incidents per module $m = (c, k)$ and the total number of incidents per cyber risk category c, respectively.

Poisson Process A simple counting process for incidents – reflecting non-stationarity of cyber risk – is a *time-inhomogeneous* Poisson process with intensity function λ^m per firm for cyber risk module m.

Definition 2.1 (Time-inhomogeneous Poisson process)**.** *A counting process $(\mathcal{N}_t)_{t\geq 0}$ is called a time-inhomogeneous Poisson process on $(\Omega, \mathcal{F}, \mathbb{P})$ with locally integrable rate (or intensity) function $\lambda : [0, \infty) \to [0, \infty)$ if:*

1. *$\mathcal{N}_0 = 0$,*
2. *the process has independent increments,*
3. *for any time interval $(s, t]$, the number of incidents is Poisson distributed with mean $\int_s^t \lambda(u)\, du$, i.e.,*

$$\mathcal{N}_t - \mathcal{N}_s \sim \text{Poiss}\left(\int_s^t \lambda(u)\, du\right).$$

Unless the intensity function is constant, the increments of a time-inhomogeneous Poisson process are *non-stationary*. The cumulative rate function $\int_0^t \lambda(u)\, du$ corresponds to the expected number of incidents up to time t.

Zeller and Scherer (2021) adopt this approach for idiosyncratic incidents. For each policyholder i of group k and module $m = (c, k)$, the number of idiosyncratic incidents $(\mathcal{N}_t^{m,i})_{t\geq 0}$ is assumed to follow a time-inhomogeneous Poisson process with intensity $\lambda^m = \lambda^{(c,k)}$. Clearly, for each cyber risk category c, the intensity depends on the covariates x^k of group k (but not on the individual policyholder i), and Zeller and Scherer (2021) propose a *generalized additive model*

$$\lambda^{(c,k)}(t) = \exp(f^c(x^k) + g^c(t))$$

to estimate the intensity rates.[3] In particular, similarities and deviations of the risk profiles of the K groups – expressed in terms of the covariate vectors x^k, $k = 1, \ldots, K$ – are reflected by the intensity functions $\lambda^{(c,k)}$.

[3]The auxiliary function f additively maps the covariates, while g captures the time dependence.

Since idiosyncratic incidents are independent across firms, the total number of incidents $\mathcal{N}_t^{m,agg}$, $t \geq 0$, per module $m = (c,k)$ as well as the total number of incidents $\mathcal{N}_t^{(c)}$, $t \geq 0$, per cyber risk category c, respectively, are again time-homogeneous Poisson processes with respective intensities

$$\lambda^{m,agg}(t) = n_k \lambda^{(c,k)}(t), \quad \lambda^{(c)}(t) = \sum_{k=1}^{K} n_k \lambda^{(c,k)}(t), \quad t \geq 0. \tag{1}$$

Cox Process More delicate, however, is the case of systematic cyber risk incidents. In particular, frequency distributions of different policyholders might be subject to dependencies due to joint underlying cyber risk factors $\mathcal{R}_1, \ldots, \mathcal{R}_d$, representing for example commonly used software that could be exploited. Such dependencies between counting processes can be captured in the context of *Cox processes*, also called *doubly stochastic Poisson processes*, extending the notion of a time-inhomogeneous Poisson process to a random intensity.

Definition 2.2 (Cox process)**.** *A Cox process $(\mathcal{N}_t)_{t\geq 0}$ is a counting process described by a random intensity process $(\lambda_t)_{t\geq 0}$ such that conditional on the specific realization $t \mapsto \lambda_t(\omega)$, $\omega \in \Omega$, the process $(\mathcal{N}_t)_{t\geq 0}$ is a time-inhomogeneous Poisson process with intensity $t \mapsto \lambda(t) = \lambda_t(\omega)$.*

A reasonable assumption could be that the intensity is a function of the current state of the cyber risk factors, i.e., for an $\mathbb{R}^d$-valued stochastic process $\mathcal{R}_t = (\mathcal{R}_t^1, \ldots, \mathcal{R}_t^d)$, $t \geq 0$, of cyber risk factors and a function $\lambda : \mathbb{R}^d \to [0, \infty)$, the intensity process is defined as

$$\lambda_t(\omega) = \lambda(\mathcal{R}_t(\omega)), \quad t \geq 0, \, \omega \in \Omega.$$

More generally, the intensity process could be modeled as a function of the whole history of cyber risk factors, i.e.,

$$\lambda_t(\omega) = \lambda(\mathcal{R}_u(\omega) : u \leq t), \quad t \geq 0, \, \omega \in \Omega.$$

In sum, in the case of systematic cyber risk, a reasonable model for the number of incidents $\mathcal{N}_t^{m,i}$ up to time t per module m and policyholder i of group k could be assuming that $(\mathcal{N}_t^{m,i})_{t\geq 0}$ follows a Cox process with intensity process $\lambda_t^m = \lambda^m(\mathcal{R}_t)$, $t \geq 0$, defined in terms of a suitable function $\lambda^m : \mathbb{R}^d \to \infty$, such that conditional on the cyber risk factors $t \mapsto \mathcal{R}_t(\omega) = (\mathcal{R}_t^1(\omega), \ldots, \mathcal{R}_t^d(\omega))$ the counting processes $(\mathcal{N}_t^{m,i})_{t\geq 0}$, $m = (c,k)$, $c = 1, \ldots, C$, $k = 1, \ldots, K$, are independent time-inhomogeneous Poisson processes. In particular, conditional independence implies that – conditional on the specific realization $t \mapsto \lambda_t^m(\omega)$ – the total number of incidents $\mathcal{N}_t^{m,agg}$, $t \geq 0$, per module $m = (c,k)$ and the total number of incidents $\mathcal{N}_t^{(c)}$, $t \geq 0$, per cyber risk category c are again time-inhomogeneous Poisson processes with intensities

$$\lambda_t^{m,agg} = n_k \lambda_t^{(c,k)}, \quad \lambda_t^{(c)} = \sum_{k=1}^{K} n_k \lambda_t^{(c,k)}, \quad t \geq 0,$$

in analogy to (1).

In contrast to the time-inhomogeneous Poisson process, the increments of a Cox process $(\mathcal{N}_t)_{t\geq 0}$ are in general no longer independent, but subject to autocorrelation. More precisely, for any $s < t \leq u < v$, the tower property of conditional expectation implies

$$\operatorname{Cov}(\mathcal{N}_t - \mathcal{N}_s, \mathcal{N}_v - \mathcal{N}_u) = \operatorname{Cov}\left(\int_s^t \lambda_z \, dz, \int_u^v \lambda_z \, dz \right),$$

i.e., the autocorrelation depends on the random intensity process. Statistical analyses of Bessy-Roland, Boumezoued, and Hillairet (2021) yield empirical evidence for autocorrelation in the number of attacks, and thus provide an additional reasoning for using Cox processes when modeling claims frequency. However, the specification of the intensity process to reproduce empirical autocorrelation is challenging.

Hawkes Process To cope with the stylized fact of autocorrelation between the number of cyber attacks, Bessy-Roland, Boumezoued, and Hillairet (2021) focus on the class of self-exciting *Hawkes processes.*

Definition 2.3 (Hawkes process)**.** *A one-dimensional Hawkes process $(\mathcal{N}_t)_{t\geq 0}$ is a point process with jump times $T_1, T_2, \ldots$ and with random intensity $t \mapsto \lambda_t$, given by*

$$\lambda_t = \mu(t) + \sum_{T_n \leq t} \varphi(t - T_n) = \mu(t) + \int_{[0,t)} \varphi(t-u)\, d\mathcal{N}_u,$$

where $\mu(\cdot)$ is a baseline intensity of jumps, and where φ is the excitation function or kernel function resp. which expresses the positive influence of past incidents at time T_n on the current value of the intensity.

From a conceptual point of view, Hawkes processes allow to capture – besides autocorrelation of the number of cyber risk incidents – excitation effects, by coupling the arrival rate of events with the number of past incidents. In particular, this allows modeling systematic incidents that affect a very large number of counterparties at the same time, e.g., exploits of widely used software such as Windows or MacOS.

Self-excitation of cyber incidents for each policyholder as well as the excitation between policyholders of different groups can be modeled by a multivariate Hawkes model. More precisely, for all cyber risk modules $m = (c, k)$ and for any policyholder i of group k, the intensity of the counting process $(\mathcal{N}_t^{m,i})_{t\geq 0}$ takes the form

$$\lambda_t^{(c,k,i)} = \mu^{(c,k)}(t) + \sum_{l=1}^{K} \sum_{j=1}^{n_l} \sum_{T_n^{(c,l,j)} \leq t} \varphi_{i,j}^{c,k,l}(t - T_n^{(c,l,j)}),$$

where

- $t \mapsto \mu^{(c,k)}(t)$ is the deterministic base intensity function, depending on the cyber risk module $m = (c, k)$ only,
- $t \mapsto \varphi_{i,j}^{c,k,l}(t)$ are self- and mutually-exciting maps (called kernels), depending on both the cyber risk module $m = (c, k)$, the other group l and the individual policyholders i, j,
- and $T_n^{(c,l,j)}$, $n \in \mathbb{N}$, are the claims arrival times of policyholder j in group l with respect to the cyber risk category c.

In this multivariate Hawkes model, the kernels $\varphi_{i,i}^{c,k,k}$ describe the self-excitation for policyholder i of group k, while the $\varphi_{i,j}^{c,k,l}$ for different policyholders $i \neq j$ model contagion between policyholders and across groups.

Using suitable parametric functions for the kernels $\varphi_{i,j}^{c,k,l}$, both the baseline intensity and the kernels can in principle be estimated by the Maximum-Likelihood method – provided that data is available in the desired amount and granularity.

2.1.2. Severity

Every claim occurring in the frequency-severity model triggers a loss size that is modeled as a random variable. The key governing parameter for the choice of the claim size distribution is the incident category c; characteristics of group k then determine distributional details, e.g., parameter values. Let $\mathcal{Y}_j^{m,i}$ denote the claim size of the jth event allocated to firm i for module

$m = (c, k)$. We assume that $(\mathcal{Y}_j^{m,i})_{j\geq 1}$, $i = 1, \ldots, n_k$, is a collection of non-negative independent[4] and identically distributed random variables.

Due to the limited availability of loss data, empirical research on cyber risk severity distributions has mostly focused on the category of data breaches. For this category, open source data bases, such as the Privacy Rights Clearinghouse Chronology of Data Breaches, are available and regularly updated. Data breach severities are found to follow strongly heavy-tailed distributions such as power-law (see, e.g., Maillart and Sornette (2010)), log-normal (see, e.g., Edwards, Hofmeyr, and Forrest (2016)) or generalized Pareto distributions (GPD) (see, e.g., Wheatley, Maillart, and Sornette (2016) or Sun, Xu, and Zhao (2020)). For cyber risk categories different from data breaches, less data is publicly available. Consequently, fewer studies have appeared that empirically analyzed the respective severity distributions. A noticeable exception are analyses based on operational risk data bases such as Biener, Eling, and Wirfs (2015) or Eling and Wirfs (2019). These approaches possess the advantage of studying all categories of cyber incidents simultaneously. In particular, Eling and Wirfs (2019) detect distributional differences between small and large claim size distributions for all considered cyber incident categories. The authors thus propose a *composite distribution approach*, where excess losses over a threshold are modeled using a GPD and the remaining smaller losses are modeled using a simple parametric distribution such as a gamma or log-normal distribution. In general, composite distribution approaches constitute a flexible modeling tool to take the empirically observed distributional differences between body and tail of severity distributions adequately into account. A composite distribution approach can be formalized as follows.

For each module m, we choose a threshold θ^m distinguishing small from large cyber claims. Small and large claims, i.e., the body and tail of the severity distribution, are then modeled separately: The i.i.d. claim sizes follow a composite distribution with density

$$f_{\mathcal{Y}_j^m}(y) := \begin{cases} C^m \cdot f_{\text{small}}^m(y), & \text{if } -\infty < y \leq \theta^m, \\ C^m \cdot f_{\text{large}}^m(y), & \text{if } \theta^m < y < \infty, \end{cases}$$

where $f_{\text{small}}^m, f_{\text{large}}^m$ are probability density functions modeling the sizes of small and large claims in module m, respectively, and C^m is a normalizing constant resulting from continuity (and possibly also differentiability) conditions at the threshold θ^m. Depending on the characteristics of the module m, different choices for $f_{\text{small}}^m, f_{\text{large}}^m$ may be suitable. Examples include

- **Small Claims:** PERT, Normal, Gamma, Log-Normal, Kernel Distribution, GPD
- **Large Claims:** GPD

The composite distribution approach is well-suited for modeling non-life insurance severity distributions in general, and cyber risks in particular.[5] Note, however, that it is independent of time, i.e., it provides only a snapshot of the current cyber environment. Due to the fast-evolving, non-stationary cyber landscape, the suitability of the model needs to be regularly checked and updated. For further details as well as summaries of other approaches to cyber severity analyses, we refer the interested reader to the excellent summaries provided by Zeller and Scherer (2021), Section 2.1, or Eling (2020), in particular Table 4 and 6, and to Cooray and Ananda (2005) for the introduction and an application of composite distributions in a non-cyber specific context.

[4] Cyber event claim sizes in a certain time interval may not always be independent; e.g., due to commonly used cyber security measures. The resulting dependence structures could be captured by alternatively imposing *conditional* independence assumptions given a set of joint underlying risk factors – similar to the idea of Cox processes as described above.

[5] Sun, Xu, and Zhao (2020) also suggest a composite distribution approach for modeling malicious hacking data breach risk. Here, the tail of the distribution follows a GPD, and the distribution body is modeled using a non-parametric kernel distribution. Due to both its suitability to and flexibility, a similar approach is also incorporated in the cyber risk model of Zeller and Scherer (2021).

2.2. Dependence Modeling

The distribution of the total claims amount per module and at portfolio level is affected by the underlying dependence structures. For cyber risk, dependencies are present in different ways including:

- dependence between frequency and severity of a certain cyber risk (category) distribution, in contrast to the classical framework of frequency severity models (e.g., due to a low level of cyber self protection of a firm),
- dependence between frequency distributions of different policyholders (e.g., due to commonly used software, such as Windows or MacOS, that could be exploited),
- dependence between severity distributions of different policyholders (e.g., due to commonly used IT security measures).

In this section, we analyze classical concepts of modeling such dependencies in a cyber risk context.

2.2.1. Common Risk Factors and Correlation

Two of the most classical concepts used in cyber risk dependence modeling are (linear) correlation and common risk factors. The (linear) correlation coefficient

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} \in [-1,1]$$

captures a possible linear relationship between the random variables X and Y: The extreme values of 1 and -1 indicate bivariate distributions entirely supported on an upward or downward sloping line, respectively. Thus, linear correlation captures dependence on a macro-level. Common risk factors, in contrast, model dependence for systematic risks on a micro-level. Here, the main idea is similar to the concept behind Cox processes: There exist explicitly modeled underlying risk factors to which all risks are jointly exposed.

Both of these classical concepts have been widely used in the cyber risk modeling literature. For example, they are key elements of the cyber risk models proposed by Böhme (2005), Zeller and Scherer (2021), and Böhme and Kataria (2006). Böhme (2005) models dependence using one common risk factor. This factor represents a common vulnerability in a portfolio of n individual risks. The connection between individual risks and the latent risk factor is then modeled and studied using linear correlation. More recently, common risk factors have also appeared in the cyber risk model of Zeller and Scherer (2021). Here, the authors use marked point processes with two-dimensional marks: the first component describes the strength of an attack, and the second component represents the subset of companies affected. Dependence among firms occurs due to the restriction of incidents to certain industry sectors. This sector-specificity is modeled through a common Bernoulli risk factor. In Böhme and Kataria (2006), the authors observe that different types of cyber incidents such as hacker attacks, hardware failures, viruses or phishing attacks vary in the type of induced intra- and inter-firm dependence. Hence, Böhme and Kataria (2006) propose a "twin-tier"-approach modeling internal (within a firm) and global linear correlation (across firms) separately. To be precise, Böhme and Kataria (2006) suggest beta-binomial marginal distributions adapted to the chosen internal correlation ρ_I and assume a certain dependence structure (a t-copula) between the margins with given global correlation ρ_G.

A given linear correlation coefficient does not fully determine the dependence structure between random variables. Many dependence structures exist that are consistent with the same linear correlation, but that may strongly differ from each other. The dependence structure of the components of a random vector is fully determined by their copula.

2.2.2. Copulas

In actuarial applications, copulas are a standard tool for modeling dependencies. A d-dimensional *copula* $\mathcal{C} : [0,1]^d \to [0,1]$ is the distribution function of a d-dimensional random vector with uniform one-dimensional marginal distributions unif$([0,1])$. Copulas provide a detailed understanding of dependence beyond linear correlation.

Theorem 2.4 (Sklar's Theorem)**.** *1. For any d-dimensional distribution function F with margins $F_1, \ldots, F_d$ there exists a copula $\mathcal{C}$ with*

$$F(x_1, \ldots, x_d) = \mathcal{C}(F_1(x_1), \ldots, F_d(x_d)) \quad \text{for all } x_1, \ldots, x_d \in [-\infty, \infty]. \tag{2}$$

In particular, if all F_i are continuous, then $\mathcal{C}$ is unique.

2. Conversely, for a given copula $\mathcal{C}$ and given one-dimensional distribution functions $F_1, \ldots, F_d$, the function F in (2) is a d-dimensional distribution function with copula $\mathcal{C}$ and marginal distribution functions $F_1, \ldots, F_d$.

Property 1 states that a copula extracts the dependence structure of a random vector from its multivariate distribution, while property 2 provides a flexible construction principle of multivariate models by combing marginal distributions and copulas to multivariate distributions. Prominent examples for copulas include:

- **Gaussian copula:** Let Φ^{-1} be the quantile function of the standard normal distribution and Φ_Σ the joint cumulative distribution of a multivariate normal distribution with covariance matrix Σ. Then the Gaussian copula is given by

$$\mathcal{C}_\Sigma^{\text{Ga}}(u_1, \ldots, u_n) = \Phi_\Sigma(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)).$$

- **t-copula:** Let $t_{\nu,\Sigma}$ signify the distribution function of a d-dimensional t-distribution $t_d(\nu, 0, \Sigma)$ for a given correlation matrix Σ and with ν degrees of freedom, and let t_ν denote the distribution function of a univariate standard t-distribution. Then the t-copula takes the form

$$\mathcal{C}_{\nu,\Sigma}^t(u_1, \ldots, u_d) = t_{\nu,\Sigma}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d)) \quad (u_1, \ldots, u_d \in [0,1]).$$

- **Archimedean copulas**: Consider a continuous function $\psi : [0, \infty) \to [0,1]$ with $\psi(0) = 1$, $\lim_{x\to\infty} \psi(x) = 0$, and ψ strictly decreasing on $[0, \psi^{-1}(0)]$, where ψ^{-1} denotes its generalized inverse. The Archimedean copula with generator ψ is given by

$$\mathcal{C}_\psi^{\text{Ar}}(u_1, \ldots, u_n) = \psi^{-1}(\psi(u_1) + \ldots + \psi(u_n)).$$

  In particular, we obtain the so-called *Gumbel copula* for $\psi_\theta(s) = (-\ln(s))^\theta$, $\theta \in [1, \infty)$.

Copulas capture dependence on a global scale – and can thus represent various kinds of dependencies when aggregating cyber risks. For example, Herath and Herath (2011) model the loss distribution *at a single firm* using a copula that captures the dependence structure between the number of affected computers of the firm and the overall severity of the loss. Dependence *between different firms* is captured using a t-copula with a given linear correlation coefficient in Böhme and Kataria (2006).

For the particular incident category c of hacking data breaches, Sun, Xu, and Zhao (2020) model the dependence *between frequency and severity* for a single firm i in module m up to time t using a Gumbel copula, i.e., their idea corresponds to the joint distribution function

$$F_{\mathcal{N}_t^{m,i}, \mathcal{Y}_j^{m,i}}(n, y) = \exp(-[(-\ln(F_{\mathcal{N}_t^{m,i}}(n)))^\theta + (-\ln(F_{\mathcal{Y}_j^{m,i}}(y)))^\theta]^{1/\theta}), \qquad \theta \geq 1,$$

for all $j = 1, \ldots, \mathcal{N}_t^{m,i}$. In contrast to the standard independence assumption in the collective risk model, the Gumbel copula is asymmetric and thus allows to capture the statistically observed upper tail dependence for frequency and severity of hacking breaches.

Note, however, that fixed copulas cannot represent the dynamic interactions occurring during systemic incidents such as cyber epidemics on the firm level. To model this interdependence explicitly, we suggest the use of epidemic network models, as detailed in the subsequent section. Still, as we will see, copulas also appear as parts of some of these models, e.g., of systemic non-Markovian network models.

3. Systemic Cyber Risks

Systemic risk generally refers to the possibility that distortions in a system may spread across many entities and be augmented due to local or global feedback effects. It is often associated with a cascading propagation of losses such that multiple entities in a system are seriously affected within a specific period of time. In the context of cyber risks, the following definition was given by the World Economic Forum (see WEF (2016)):

> "Systemic cyber risk is the risk that a cyber event [...] at an individual component of a critical infrastructure ecosystem will cause significant delay, denial, breakdown, disruption or loss, such that services are impacted not only in the originating component but consequences also cascade into related ecosystem components [...]"

In this section, we study systemic cyber risk *models*. We focus on two different approaches: network models that capture *interconnectedness and cascading propagation* (Section 3.1), and game-theoretic models that represent the *strategic interaction* of cyber ecosystem components (Section 3.2). Pricing of systemic (and non-systemic) cyber risks is considered in Section 4.

3.1. Interconnectedness and Contagious Risk Modeling in Networks

Interconnectedness constitutes a key characteristic of cyber systems. It may trigger and amplify cyber events. Cyber network models for contagious risk propagation consist of the following three key components:

1. A **network** (also called *graph*) whose nodes represent components or agents. These entities could be individual corporations, subsystems of computers, or single devices. The edges of the network correspond to possible transition channels, e.g., IT connections or exchange of data/computer code, see Section 3.1.1;

2. A **spread process** on the network that models the propagation of a computer virus, a Trojan, or ransomware, see Section 3.1.2;

3. A **loss model** which determines the severity of cyber events and the monetary impact on different agents in the network, see Section 3.1.3.

3.1.1. Networks

Definition 3.1 (Network)**.** *A* network *(or* graph*) G is an ordered pair of sets $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} \neq \emptyset$ is a countable set of N elements, called* nodes *(or* vertices*), and $\mathcal{E}$ is a set of pairs (i, j), $i, j \in \mathcal{V}$, of different nodes, called* edges *(or* links*). If all edges in $\mathcal{E}$ are unordered, G is called an* undirected network*. Otherwise, the network G is called* directed*.*

The network structure is encoded in its *adjacency matrix* $A = (a_{ij})_{i,j\in\{1,\dots,N\}} \in \{0,1\}^{N\times N}$, which is defined by its entries

$$a_{ij} := \begin{cases} 1, & \text{if } (i,j) \in \mathcal{E} \\ 0, & \text{if } (i,j) \notin \mathcal{E}. \end{cases}$$

By definition, G is undirected if and only if A is symmetric. Examples of undirected network topologies with $N = 8$ nodes are depicted in Figure 1.

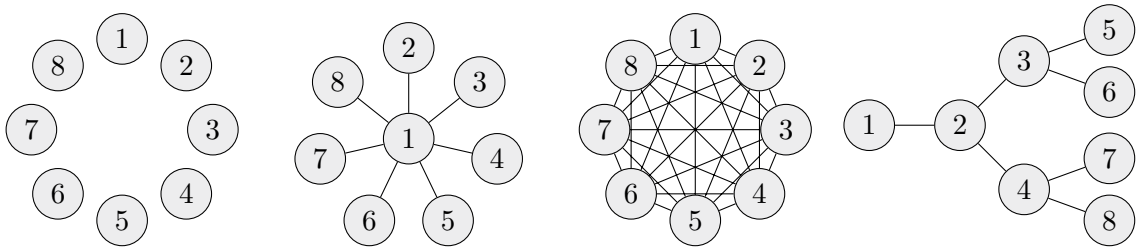


Figure 1: Examples of network topologies with $N = 8$ nodes.

In applied network analysis, the exact network structure is often unknown. In this case, random network models enable sampling from a class of networks with given fixed topological characteristics (such as the overall number of nodes). In a random network, each possible edge in the network is present (or absent) with a given fixed probability.[6]

In the cyber insurance literature, network models are mainly applied to the study of risk contagion, e.g., modeling the propagation of malware in IT networks of interconnected firms or devices. In addition to an underlying network, an adequate model of the contagion process is required that captures the epidemic spread.

[6]Two standard classes of undirected random networks are Erdős–Rényi networks and scale-free networks:

- **Erdős–Rényi networks:** The simplest random network model was introduced by Erdős and Rényi (1959): The Erdős–Rényi network $G_p(N)$ is constructed from a set of N nodes in which each of the possible $N(N-1)/2$ edges is independently present with the same probability p. The resulting degree distribution, i.e., the distribution of the number of neighbors of any node in the network, is binomial, since the probability to create a node of degree k (i.e., with k neighbors) $P(k)$ is equal to the probability that this node is connected to exactly k other nodes and not connected to the remaining $N-1-k$ nodes of the network:

  $$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}.$$

  For large N and in the limit of constant average degree $(N-1)p \approx Np =: c$, the binomial distribution can be approximated by a Poisson distribution

  $$P(k) = e^{-c}\frac{c^k}{k!}.$$

- **Scale-free networks:** Empirical analysis in various research areas suggests that real-world networks exhibit much more heterogeneous degrees than Poisson distributions would suggest. Often a hierarchy of nodes is observable – with a few nodes of high degree (called *hubs*), and a vast majority of less connected nodes having a relatively low degree. Typically, the degree distribution is approximately *scale-free*, i.e., we have

  $$P(k) \approx ak^{-\lambda}, \qquad a > 0, \quad \lambda > 0.$$

  A special case with $\lambda = 3$ is given by the *Barabási-Albert model* where a growing network is generated following a preferential attachment rule, see Barabási and Albert (1999) for details.

3.1.2. Epidemic Spread Processes

Models of infectious spread dynamics have intensively been investigated in the field of mathematical biology and epidemiology, dating back at least to the seminal work of Kermack and McKendrick (1927).[7] In this paper, we focus on epidemic *network* models for populations of entities.

At each point in time, each node is in a particular state that may change over time due to its interaction with other nodes. According to their state, individuals are divided into distinct *compartments*, e.g., individuals that are *susceptible* (S) to an infection, *infected* (I) individuals, or individuals who have *recovered* (R) from the infection. For a network of N nodes, the spread process can at each point in time t be described by a *state vector*

$$X(t) = (X_1(t), \ldots, X_N(t)) \in E^N,$$

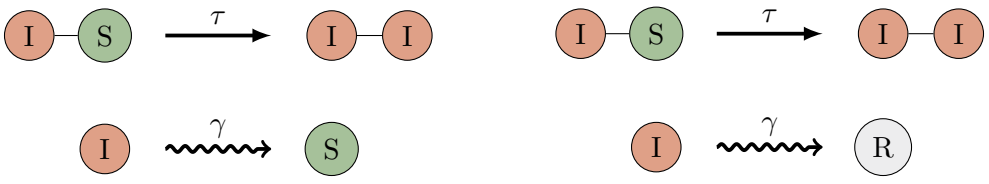
where E is the set of compartments. Both *Markov* and *non-Markov* processes have been considered in the context of epidemic spread processes.[8]

Markovian Spread Models In Markovian spread models on networks, the evolution of the state vector $X(t)$ is described by a (time-homogeneous) continuous-time Markov chain on the discrete state space E^N. The SIS *(Susceptible-Infected-Susceptible)* and SIR *(Susceptible-Infected-Recovered)* Markov models constitute the most frequently used epidemic spreading models on networks. They differ in the presence (SIR) or absence (SIS) of immunity: Reinfection events are only possible in the SIS framework, since in the SIR model, recovered individuals gain (permanent) immunity, i.e., the models build on the two different compartment sets $E = \{S, I\}$ and $E = \{S, I, R\}$, respectively.

In both models, a transition of X from one state in E^N to another is only possible if exactly one node changes its state X_i in E. State changes can occur through infection or recovery: It is assumed that each node may be infected by its infected neighbors, but can be cured independently of all other nodes in the network. Each node is endowed with an independent exponential clock and changes its state when the exponential clock rings. Letting $\tau > 0$ and $\gamma > 0$, the rates of these transitions are illustrated in Figure 2 and given as follows ($i = 1, \ldots, N$):

$$\begin{aligned} X_i : S \to I \quad &\text{with rate} \quad \tau \sum_{j=1}^{N} a_{ij} \mathbb{1}_{\{X_j(t)=I\}} \\ X_i : I \to Z \quad &\text{with rate} \quad \gamma, \end{aligned} \tag{3}$$

where $Z = S$, for the SIS, and $Z = R$ for the SIR model, respectively.



(a) SIS Model **(b) SIR Model**

Figure 2: Infection and recovery for the SIS and SIR network model.

[7] The models focus either on an epidemic spread within a population, as, e.g., in Kermack and McKendrick (1927), or on the spread along paths of a predefined network; for a detailed overview, see, e.g., Pastor-Satorras et al. (2015) and Kiss, Miller, and Simon (2017).

[8] Intuitively, the Markov property implies that a process is "memoryless", i.e., that the conditional distribution of future values X_{t+s}, $s > 0$, of the process does only depend on the present value of the process X_t and not on past values X_μ, $\mu < t$.

The exponential transition times enable an intuitive stochastic simulation algorithm: the well-known *Gillespie algorithm*, first introduced in Gillespie (1976) and Gillespie (1977).

Algorithm 3.2 (Gillespie)**.** Input: *Initial state of the system* $x_0 \in E^N$ *and initial time* $t_0 \geq 0$.

1. (Initialization) *Set the current state* $x \to x_0$ *and current time* $t \to t_0$.
2. (Rate Calculation) *For the current state of the system* x*, calculate the sum of rates for all possible transitions* $q_x = \sum_{i=1}^{N} q_{x_i}$*, where* q_{x_i} *denotes the rate for a state change of node* i *according to* (3).
3. (Generate Next Event Time) *Sample the next event time* t_{new} *from an exponential distribution with parameter* q_x.
4. (Choose Next Event) *Sample the node* i_{new} *at which the next transition occurs: Each node* $i = 1, \ldots, N$ *is chosen with probability* q_{x_i}/q_x.
   Change the state $x_{i_{new}} \to y_{i_{new}}$ *according to* (3).
5. *Set* $t \to t + t_{new}$*,* $x \to (x_1, \ldots, x_{i_{new}-1}, y_{i_{new}}, x_{i_{new}+1}, \ldots, x_N)$ *and return to Step 2 until a prespecified stopping criterion is met.*

For practical purposes such as the pricing of cyber insurance contracts, we often do not need the full information provided by the Markov chain evolution, but only the dynamics of specific quantities such as moments or (infection) probabilities. Of particular interest are the dynamics of the state probabilities of individual nodes $\mathbb{P}(X_i(t) = x_i),\ t \geq 0$. They can be derived from Kolmogorov's forward equation and written in general form as $(i = 1, \ldots, N)$

$$\frac{d\mathbb{P}(X_i(t) = x_i)}{dt} = \sum_{y: y_i = x_i} \sum_{z \neq y} [\mathbb{P}(X(t) = z) q_{zy} - \mathbb{P}(X(t) = y) q_{yz}], \tag{4}$$

where q_{zy} denotes the transition rate of the entire process X from $z \to y$. In natural sciences, this equation is also known under the term *master equation*. For the SIS and SIR models, using Bernoulli random variables $S_i(t) := \mathbb{1}_{\{X_i(t)=S\}}$, $I_i(t) := \mathbb{1}_{\{X_i(t)=I\}}$, and (for SIR) $R_i(t) := \mathbb{1}_{\{X_i(t)=R\}}$, the dynamics of state probabilities of individual nodes (4) can conveniently be written via moments:

- **SIS model:**[9] Since $E = \{I, S\}$, we have $S_i(t) = 1 - I_i(t)$, i.e., the evolution of X is fully described by the evolution of the vector $I(t) = (I_1(t), \ldots, I_N(t))$, and the single node infection dynamics are given by

$$\frac{d\mathbb{E}[I_i(t)]}{dt} = -\gamma \mathbb{E}[I_i(t)] + \tau \sum_{j=1}^{N} a_{ij} \mathbb{E}[I_j(t)] - \tau \sum_{j=1}^{N} a_{ij} \mathbb{E}[I_i(t) I_j(t)], \quad i = 1, \ldots, N, \tag{5}$$

  since $\mathbb{P}(X_i(t) = I) = \mathbb{P}(I_i(t) = 1) = \mathbb{E}[I_i(t))]$. Note that this system of N equations is not closed as second order moments $\mathbb{E}[I_i(t)I_j(t)]$, i.e., second order infection probabilities, appear.

- **SIR model:** The dynamics of the recovery Bernoulli random variable $R_i(t)$ result from the dynamics of $I_i(t)$ and $S_i(t)$ due to $\mathbb{E}[R_i(t)] = 1 - \mathbb{E}[S_i(t)] - \mathbb{E}[I_i(t)]$. Equation (4)

[9] In the cyber insurance literature, the SIS Markov model was used by Fahrenwaldt, Weber, and Weske (2018). Also, a brief application was studied in Xu and Hua (2019) with a modified ε-SIS model, originally proposed in Mieghem and Cator (2012). Here, an infectious threat for node i from outside the network is included with a rate ε_i.

corresponds to:

$$\begin{aligned}\frac{d\mathbb{E}[S_i(t)]}{dt} &= -\tau \sum_{j=1}^{N} a_{ij}\mathbb{E}[S_i(t)I_i(t)],\\ \frac{d\mathbb{E}[I_i(t)]}{dt} &= \tau \sum_{j=1}^{N} a_{ij}\mathbb{E}[S_i(t)I_j(t)] - \gamma\mathbb{E}[I_i(t)],\end{aligned} \tag{6}$$

for $i = 1, 2, \ldots, N$. Again, the system is not closed due to the presence of second order moments.

The main problem with systems (5) and (6) is the fact that they are *not closed*: They depend on second order moments, which, in turn, depend on third order moments, etc. For example, the fully closed SIS model yields $\sum_{i=1}^{N} \binom{N}{i} = 2^N - 1$ moment (i.e., infection probability) equations. Solving these systems exactly becomes intractable for networks of realistic size. To deal with this issue, the following two approximation approaches have been proposed:

1. **Monte Carlo simulation:** Monte Carlo simulation using the Gillespie algorithm (Algorithm 3.2 above) constitutes a powerful tool to obtain various quantity estimates related to the evolution of the epidemic spread. In particular, this includes the state probability dynamics of individual nodes (4).[10]

2. **Moment closures:** If a set of nodes $J \subset \mathcal{V}$ is infected, this increases the probability of other nodes in the network (that are connected to the set J via an existing path) to become infected as well. Hence, node states are to some extent *correlated*. To break the cascade of equations and to make ODE systems tractable, the moment closure approach consists in assuming independence at a certain order k, neglecting any further correlations. This is done by considering the exact moment equations up to this order k and *closing* the system by approximating moments of order $k+1$ in terms of products of lower-order moments using a mean-field function. A detailed explanation of two different types of moment closures is provided in Appendix B.

   However, a major problem with moment closures is that only little is known about rigorous error estimates.[11] This presents an important avenue for future research.

Non-Markovian Spread Models Non-Markovian models possess conditional distributions that may depend on the past and further random factors. In contrast to the Markovian setup, where transition times are necessarily exponential, non-Markovian models typically allow the flexibility to freely choose the distributions of infection and recovery times. In addition, dependence among the infection times may be included. This generality may certainly improve the quality of a fit to real-world data. However, the extended generality in comparison to Markov models is typically associated with reduced tractability. For this reason, non-Markovian are less commonly considered.

A simple example of a non-Markovian model for the spread of cyber risks has been proposed by Xu and Hua (2019). The model does not include immunity, i.e., the underlying compartment set is the same as for the Markovian SIS model. Waiting times considered in the model are

- The individual **recovery times** T_i^{recov} of *infected* nodes.

[10] Pseudocode and further explanations of the Gillespie algorithm applied to the SIS and SIR epidemic network models is, e.g., given in Appendix A.1.1 of Kiss, Miller, and Simon (2017).

[11] This problem has also been highlighted in the epidemic literature, see, e.g., Kiss, Miller, and Simon (2017), p.115.

- For nodes i which are in the *susceptible* state, two different types of infections are considered, *internal infections* from within the network and *external infections* comming from outside:
  1. **Internal infection times:** Let the random variable $K_i(t) = \sum_{j=1}^{N} a_{ij} I_j(t)$ denote the number of infected neighbors of node i at time t. Infectious transmissions to node i are given with waiting times $T_{i_1}, \ldots, T_{i_{K_i}}$. These times share the same marginal distribution F_i. Their underlying dependence structure is captured by a prespecified copula.
  2. **External infection times:** A random variable T_i^{out} with distribution G_i models the arrival time of threats from outside the network to node i. T_i^{out} is assumed to be independent of times $T_{i_1}, \ldots, T_{i_{K_i}}$.

To simulate the process, the waiting times for all nodes are generated according to their current state (i.e., recovery times for all infected nodes, and internal and external infection times for all infected nodes). The minimum of these waiting times determines the next event (infection or recovery). After this change, the process is repeated until a prespecified stopping criterion is met.[12]

Finally, note that a Markovian SIS model with outside infections[13] can be obtained as a special case by choosing exponentially distributed infection and recovery times and assuming independence between all waiting times.

Top-Down Approaches The spread models proposed so far are bottom-up models where exact infection dynamics are studied. This requires complete information about the underlying network topology. A major challenge in real-world applications is the fact that the exact network structure is often unknown. This makes a more general risk assessment desirable.

For this reason, an entirely different approach was taken in Hillairet and Lopez (2021). The authors determine the impact of massive global-scale cyber-incidents, like the WannaCry scenario, on insurance losses and assistance services. While network contagion is implicitly considered, it is not modeled within an actual network framework; instead, the authors choose the original population-based SIR model of Kermack and McKendrick (1927) to determine the *deterministic* dynamics of the total numbers of susceptible, infected, and recovered individuals within the global population of IT devices. The corresponding ODE system is given by

$$\begin{aligned} \frac{dS(t)}{dt} &= -\tau S(t) I(t) \\ \frac{dI(t)}{dt} &= \tau S(t) I(t) - \gamma I(t) \\ \frac{dR(t)}{dt} &= \gamma I(t) \end{aligned}$$

with constant global population size $N = S(t) + I(t) + R(t)$.

Given this global spread, the focus of the paper lies on the *stochastic* evolution of the insurer's *local* portfolio consisting of $n << N$ policyholders and their corresponding losses. The influence of the global cyber epidemic on the local portfolio is captured by the hazard rate $\lambda_{T_i^{infec}}$ of the policyholders' infection times T_i^{infec}:

$$\lambda_{T_i^{infec}}(t) = \lim_{dt \to 0^+} \frac{1}{dt} \mathbb{P}(T_i^{infec} \in [t, t+dt] \mid T_i^{infec} \geq t) := \tau I(t),$$

i.e., the local hazard rates are assumed to be proportional to the number of infected individuals in the global population.

[12] Pseudocode for stochastic simulations is provided in Algorithm 1 of Xu and Hua (2019).

[13] To be precise, the so-called ε-SIS model, originally proposed in Mieghem and Cator (2012), arises.

Due to the scarcity of data currently available, such top-down approaches present promising avenues for future research on cyber risk and insurance.

3.1.3. Loss Models

Given the underlying network, and the epidemic spread process X on it, the third and final ingredient of a cyber risk network model is given by a suitable loss model $Y_{i,j}$ for each node $i = 1, \dots, N$, where j describes the number of loss events. In the existing literature, loss models are kept rather simple as the focus lies on modeling the cyber-epidemic spread. We give two examples:

1. In Fahrenwaldt, Weber, and Weske (2018), cyber attacks are launched in a two-step procedure: First, using a homogeneous Poisson process, times of attacks on the entire network (loss events) $t_1, t_2, \dots$ are generated. Second, for each node i, a possible random loss $L_{i,j}$ is modeled, where j describes the index of the corresponding attack time. The loss, however, only materializes if node i is infected at the attack time. This is captured by the loss model

$$Y_{i,j} = L_{i,j} \cdot \mathbb{1}_{X_i(t_j)=I}, \quad i = 1, \dots N, \quad j = 1, 2, \dots .$$

2. In Xu and Hua (2019), the loss model $Y_{i,j}$ is given by

$$Y_{i,j} = \eta_i(D_{i,j}) + C_i(T_{i,j}^{recov}), \quad i = 1, \dots N, \quad j = 1, \dots, M_i(T)$$

with a legal cost function η_i, the number $D_{i,j}$ of data damaged in the infection j, the total number $M_i(T)$ of infections of node i up to time T, and a cost function C_i depending on the recovery time $T_{i,j}^{recov}$ of node i for infection event j. Here, both the number of infection events $M_i(T)$ and the recovery time $T_{i,j}^{recov}$ for each event j are derived from the infection dynamics X_i up to time T while the data loss sizes $D_{i,j}$ are assumed to follow a beta distribution.

Future research should analyze the implementation of more realistic loss models, that, e.g., contain different types of cyber events and capture their characteristic severity distributions (see also the discussion on classical frequency-severity approaches in Section 2.1.2). This would strengthen the applicability of network models in real-world insurance and risk management contexts and, thus, overall help to provide a safer cyber landscape.

3.2. Game-Theoretic Models and Strategic Interaction Effects

The risk exposure of individuals is often interdependent, since it is influenced by the behavior of other actors. In addition to contagion due to the interconnectedness of entities in cyber networks, potentially different objectives of the involved actors and their strategic interaction constitute a key characteristic of systemic cyber risk. *Game theory* provides a suitable framework to study these components of risk in the cyber ecosystem.

In the first part of this section, we briefly review and provide a short mathematical introduction to game theoretic approaches applied to study cyber risk and cyber insurance (Section 3.2.1). For an exhaustive review of the corresponding literature, we refer to the surveys Böhme and Schwartz (2010), Böhme, Laube, and Riek (2018), and Marotta et al. (2017). We will adopt the notation from Marotta et al. (2017). Section 3.2.2 evaluates the considered models.

3.2.1. Game-Theoretic Modeling Approaches

The majority of game theoretic contributions focuses on self protection of interdependent actors in the presence or the absence of cyber insurance. A key question is whether and under which

conditions cyber insurance provides incentives for self protection and improves global IT security. In this section, we present[14] the main ideas and characteristics of such models.

Three Different Types of Actors in the Game We consider three types of strategic players with potentially different objectives; potential buyers of insurance, insurance companies, and the regulator:

1. **Agents** are the potential cyber insurance policyholders. To capture interdependence, most models assume that agents form a network. Agent i aims to maximize her expected utility
$$\max \mathbb{E}[U_i(W_i)],$$
where
   - U_i denotes the utility function of agent i. Various types of utility functions are considered in the literature; most of them satisfy the classical von-Neumann-Morgenstern axioms. While some papers, such as Naghizadeh and Liu (2014), Pal (2012), and Pal et al. (2014), allow for *heterogeneous preferences*, the majority of models assumes *homogeneous preferences*, i.e., $U_i = U$ across all agents.
   - W_i is the financial position of agent i at the end of the insurance period. The value W_i depends on whether the agent has bought an insurance contract or not, on her investment C_i in cyber security, and on potential losses L_i in case the agent is affected by a cyber attack.

   The agent's **self protection level** x_i is a *crucial model component* when studying interdependence.[15] Most of the existing literature falls under either of the following two distinct categories: Some assume that only two security states are possible, secured or not, with the corresponding constant cost C or 0. Others propose a continuous scale of security levels, e.g., $x_i \in [0, 1]$. The value of x_i affects
   - *the cost of self protection* C_i:
     For a continuous spectrum of security levels, i.e., $x_i \in [0, 1]$, $C_i = C(x_i)$ is typically assumed to be an increasing convex function of x_i, reflecting that user costs rapidly increase when improving security.
   - *agent i's probability of becoming infected* $p_i := \mathbb{P}(I_i = 1)$:
     Obviously, this probability depends on the individual security level x_i of the agent i, but – due to interdependence – it may also be influenced by the individual security levels of other network participants.

   Within this framework, agent i's expected utility can be computed

   a) **without insurance:**
$$\mathbb{E}[U_i(W_i)] = (1 - p_i) \cdot U_i(W_i^0 - C_i) + p_i \cdot U_i(W_i^0 - L_i - C_i)$$

   b) **with insurance:**
$$\mathbb{E}[U_i(W_i)] = (1 - p_i) \cdot U_i(W_i^0 - \pi_i - C_i) + p_i \cdot U_i(W_i^0 - L_i - C_i - \pi_i + \hat{L}_i)$$

   where
   - W_i^0 denotes the initial wealth of agent i.

[14] We refer to Marotta et al. (2017) for an in-depth overview of the topic.

[15] Only few papers, e.g. Böhme (2005), Böhme and Kataria (2006), Johnson, Laszka, and Grossklags (2014a) and Johnson, Laszka, and Grossklags (2014b), do not include self protection in the model.

- π_i is the insurance premium of agent i set by the insurer. This premium depends on the type of insurance market; we will discuss different models below.
- L_i is the potential loss of agent i that is governed by a binary distribution: only two possible scenarios are considered. Either the agent experiences a cyber attack with a *fixed loss size*, or she is not attacked which corresponds to no loss. This particular setting excludes the possibility of different types of cyber attacks. Multiple attacks are also not considered.[16] The majority of game theoretic models relies on the assumption of constant homogeneous losses for all agents, i.e., $L_i \equiv L$.
- $\hat{L}_i$ is the cover in case of loss which is specified in the insurance contract. Most papers assume full coverage, i.e., $\hat{L}_i = L_i$, but some consider alternatively partial coverage, e.g., in order to mitigate the impact of information asymmetries, cf. Mazzoccoli and Naldi (2020), Pal (2012), Pal et al. (2014).

2. **Insurance Companies**: The insurer defines cyber insurance premiums and specifies the insurance cover $\hat{L}_i$. Insurance premiums depend on the market structure:
   - **Competitive market**: This is the prevailing model in the literature. The profits of the insurers are zero in this case; customers pay fair premiums. Competitive markets are a boundary case that almost surely leads to the insurer's ruin in the long run.
   - **Monopolistic market / Representative insurer**: Another extreme is a market with only one insurance company. In these models, the impact of a monopoly can be studied. An alternative consists in studying objective functions that are different from the insurer's profit. This situation is mostly studied in the context of regulation: The insurer represents a regulatory authority and is not aiming for profit maximization, but focuses on the wealth distribution in order to incentivize a certain standard of IT protection.[17]
   - **Immature market/Oligopoly**: Instead of a monopoly, imperfect competition is studied with multiple insurers that may earn profits. The increments between the fair price and the insurance premium is determined by the markets structure.[18]

3. **Regulator:** Market inefficiencies and a lack of cyber security may be mitigated by regulatory policies. *Regulatory instruments* include mandatory insurance, fines and rebates, liability of contagion, etc. The choice of policies and their impact can be studied[19] by introducing a third party, the regulator. The objective of the regulator is maximize a *social welfare function*. This could, for example, be chosen as the sum of the expected utilities of the agents
$$\sum_i \mathbb{E}[U_i(W_i)].$$

Interdependent Self Protection in IT Networks The strategic interaction of the three types of players introduced above is modeled as a game. The agents form an interconnected network and optimize their expected utility. Their individual security level and the amount of cyber insurance coverage serve as their controls. The insurance companies are provider of risk management solutions. In some models, a regulator is included as a third party with the aim to improve welfare, e.g., by implementing standards of protection in cyber systems.

[16] We will discuss the scope of the existing models in Section 3.2.2.

[17] Market models of this type are studied in Naghizadeh and Liu (2014) with a zero-profit insurer. Profits are still possible in Pal (2012), Pal et al. (2014) and Pal et al. (2019), and maximized in Khalili, Naghizadeh, and Liu (2017).

[18] Immature markets are considered, e.g., in Martinelli et al. (2017), Martinelli and Yautsiukhin (2016), Ogut, Menon, and Raghunathan (2005).

[19] The effects of such regulatory instruments were, e.g., studied in Bolot and Lelarge (2009), Naghizadeh and Liu (2014), Pal (2012), Pal et al. (2014).

The network topologies are, typically, quite stylized to guarantee tractability. For example, two-agent models are considered in Ogut, Menon, and Raghunathan (2005). Most papers investigate complete graphs, e.g., Ogut, Menon, and Raghunathan (2005), Schwartz and Sastry (2014) and Pal et al. (2014). Bolot and Lelarge (2009) and Yang and Lui (2014), in contrast, investigate networks with degree heterogeneity, but restrict their analysis to Erdős-Rényi random graphs.

Agents are interdependent in the network, since the infection probability p_i depends on the local security level x_i and levels of the other nodes $y_i := (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N)$ (or at least of i's neighbors). In some cases, p_i is assumed to depend on an overall network security level as well.[20] However, in contrast to the models from Section 3.1, attacks do not result from a dynamic contagion process; instead, the infection is assumed to be *static* and the values p_i are derived from *ad hoc schemes*. The most common one[21] assumes a continuous spectrum of security levels and computes p_i as the complementary probability of the case that neither a direct nor an indirect attack occurs:

$$\begin{aligned} p_i(x_i, y_i) &= 1 - (1 - p_i^{dir})(1 - p_i^{cont}) \\ &= 1 - (1 - \psi_i(x_i)) \times \prod_{j \neq i} (1 - h_{i,j}\psi_j(x_j)) \end{aligned}$$

where

- $p_i^{dir} = \psi_i(x_i)$ denotes the probability of direct infection of i through threats from outside the network. It is interpreted as a function of the individual security level x_i.
- $p_i^{cont} = 1 - \prod_{j \neq i}(1 - h_{i,j}\psi_j(x_j))$ is the probability for node i to become infected through contagion. The probability for i to be infected via node j is given by $h_{i,j}$, i.e., $h_{i,j} \neq 0$ only if i and j are adjacent. This is where the underlying network topology comes into play.

In the absence of information asymmetries, the cooperative game between agents and the insurer(s) involves three perspectives:[22]

1. A legal framework is set by the regulator (if a regulator is present).
2. Agents specify their levels of self protection and insurance protection and select the available contract types to maximize their expected profits.
3. Insurance companies compute the corresponding contract details, i.e., premiums π_i and indemnities $\hat{L}_i$. In absence of information asymmetries between agents and the insurer(s), the protection levels of policyholders can be observed by the insurer and are reflected by the contract.

The model may be augmented to incorporate information asymmetries:

- **Moral hazard**: A dishonest policyholder may behave in a way that increases the risk, and the insurer cannot observe the policyholder's behavior after the conclusion of a contract. In the game, this is represented by the possibility for agents to change their self protection level after step 3.

[20] This is the case in Shetty, Schwartz, and Walrand (2010), Shetty et al. (2010) and Schwartz and Sastry (2014).

[21] An alternative approach using a simplified two-state scenario of security investments is analyzed in Bolot and Lelarge (2008a), Bolot and Lelarge (2008b), and Yang and Lui (2014). Infection probabilities are derived from a recursive branching process.

[22] Variations of the game design are possible; e.g., in Laszka, Panaousis, and Grossklags (2018) the authors use a signaling game instead of a cooperative game model to study the adverse selection problem, allowing insurers to audit the agents' security. A similar game is considered in Khalili, Naghizadeh, and Liu (2017) who introduce a pre-screening procedure.

- **Adverse selection**: Agents with larger risks have a higher demand for insurance than safer ones. The degree of the policyholders' risk tolerances cannot be observed by the insurer. The self protection levels of policyholders is not precisely known by the insurer when the contract details are computed.

In most papers, cyber insurance is not associated with additional incentives to enhance self protection. In contrast, agents may prefer to buy insurance instead of investments in self protection, i.e., from a welfare perspective, they *underinvest* in security.

These observations may be interpreted as an indication that *regulatory interventions* are necessary, such as fines and rebates, mandatory cyber insurance, or minimal investment levels.[23]

3.2.2. Evaluation of Game Theoretic Modeling Approaches

Many questions for future research remain to be answered, since the existing game theoretic models of cyber insurance and cyber security are oversimplified:

- **Simplified network topologies**: In the vast majority of the discussed literature, networks are assumed to be homogeneous. However, agents are typically heterogeneous in reality which substantially alters the cyber ecosystem. Network contagion and cyber loss accumulation are highly sensitive to the topological network arrangement; for example, important determinants are the presence (or the absence) of central hub nodes or clustering effects, see, e.g., Fahrenwaldt, Weber, and Weske (2018). For appropriate risk measurement and management these aspects need to be taken into account explicitly.
- **Static contagion**: A key feature of cyber risk in networks is the systemic amplification of disturbances. From the insurer's perspective, the contagion dynamics will clearly influence tail risks; an example are catastrophic incidents that affect a large fraction of its portfolio. Such events may be critical in terms of the insurer's solvency. An understanding of cyber losses and an evaluation of countermeasures requires dynamic models of contagion processes.
- **Constant losses**: In all considered game-theoretic models, the agent's losses are assumed to be constant, i.e., modeled as binary random variables. However, in reality we observe that the severity of instances varies substantially due to the heterogeneity of cyber events, ranging from mild losses (e.g. malfunctioning of email accounts) to very large losses (e.g. attacks on production facilities or systemic breakdowns).

Cyber insurance and instruments to control cyber risk depend on the structures of networks, the dynamics of epidemic spread processes, as well as loss models – and vice versa. These feedback loops need to be properly incorporated in future research. Key ingredients of systemic cyber risks – the interconnectedness captured by epidemic network models, and strategic interaction described in game-theoretic models – must be combined.

4. Pricing Cyber Insurance

Cyber risk comprises both *non-systemic risk*, further subdivided into *idiosyncratic* and *systematic cyber risk*, cf. Section 2, and *systemic risk*, cf. Section 3. Classical actuarial pricing, however, relies on the *principle of pooling*, and it is thus applicable for idiosyncratic cyber incidents only. For systematic and systemic cyber risk, the appropriate pricing of insurance contracts requires more sophisticated concepts and techniques.

[23] The effect of fines and rebates was studied in papers Bolot and Lelarge (2009), Naghizadeh and Liu (2014), Pal (2012), and Pal et al. (2014). In the presence of information asymmetries, fines and rebates cannot easily be applied. An alternative regulatory instrument are requirements on minimal investment levels for IT security. However, Shetty, Schwartz, and Walrand (2010), Shetty et al. (2010) and Schwartz, Shetty, and Walrand (2013) argue against such requirements.

4.1. Pricing of Non-Systemic Cyber Risks

In non-life insurance, contracts are usually signed for one year. At renewal time, the insurer may adjust premium charges as well as terms and conditions, while the policyholder can decide whether or not to continue the contract. Premium calculation thus typically refers to loss distributions on a one-year time horizon. In this section, we adopt this market convention and consider premiums payable annually in advance.[24]

As introduced in Section 2, losses and associated premiums are considered in the granularity of cyber risk categories $c \in \{1, \dots, C\}$ and homogeneous groups $k \in \{1, \dots, K\}$ of policyholders. Each pair $m = (c, k)$ is called a cyber risk module. In terms of a modular system, the premium per risk category serves as a component for the overall premium. Homogeneous groups – specified for example in terms of covariates – correspond to tariff cells, i.e., any policyholder in group k should pay the same premium $\pi^{m,\text{non-sys}}$ per risk category. We denote by n_k the number of policyholders in group k.

To decouple the pricing of idiosyncratic and systematic cyber losses, either both components must be modeled separately (see, e.g., Zeller and Scherer (2021)) or a decomposition of the total non-systemic claims amount on a one-year time horizon is needed. In the context of Section 2, this decomposition takes – in a stylized manner – the form

$$\mathcal{S}_1^{m,\text{non-sys}} = \mathcal{S}_1^{m,\text{idio}} + \mathcal{S}_1^{m,\text{systematic}},$$

where both the total idiosyncratic claims amount $\mathcal{S}_1^{m,\text{idio}}$ and the total systematic claims amount $\mathcal{S}_1^{m,\text{systematic}}$ are described by a frequency-severity approach. According to this decomposition, the premiums for idiosyncratic and systematic cyber risk $\pi^{m,\text{idio}}$ and $\pi^{m,\text{systematic}}$, respectively, are calculated separately and aggregated to form the total premium $\pi^{m,\text{non-sys}}$.

Finally, a smoothing algorithm might be helpful in order to avoid structural breaks between the premiums of risk groups with similar covariates.

Idiosyncratic Risk Idiosyncratic cyber incidents occur within each firm independently. For homogeneous groups of policyholders, defined in terms of covariates vectors x^k, $k \in \{1, \dots, K\}$, this type of cyber risk is thus subject to pooling of risk, and hence classical actuarial pricing is still applicable. More precisely, a valuation based on means with respect to the statistical or real-world measure $\mathbb{P}$ is mathematically justified by the strong law of large numbers, i.e., for each firm i with cyber risk module $m = (c, k)$ with i.i.d. annual losses $\mathcal{S}_1^{m,i,\text{idio}} \sim \mathcal{S}_1^{m,\text{idio}}$, $i = 1, \dots, n_k$, the average claims amount tends to the expected claims amount per policyholder asymptotically:

$$\lim_{n_k \uparrow \infty} \frac{1}{n_k} \sum_{i=1}^{n_k} \mathcal{S}_1^{m,i,\text{idio}} = \mathbb{E}[\mathcal{S}_1^{m,\text{idio}}] \quad \mathbb{P}\text{-a.s..}$$

This suggests that the premium per policyholder for idiosyncratic cyber incidents should – for a large number of policyholders n_k in group k – be equal to the expected claims amount, also called *net risk premium*:

$$\pi^{m,\text{idio}} = \mathbb{E}[\mathcal{S}_1^{m,\text{idio}}].$$

However, the net risk premium is not sufficient. Indeed, in a multi-period setting, ruin theory states that ruin of the insurer occurs – no matter of the initial capital – in the long run $\mathbb{P}$-a.s. if only the net risk premium is charged, see, e.g., Mikosch (2004) and the references therein. A similar result already holds in the one-period setting: Letting the number of policyholders n_k tend to infinity, the one-period loss probability

$$\mathbb{P}\Big(n_k \pi^{m,\text{idio}} - \sum_{i=1}^{n_k} \mathcal{S}_1^{m,i,\text{idio}} < 0\Big) = \mathbb{P}\Big(\sum_{i=1}^{n_k} \frac{\mathcal{S}_1^{m,i,\text{idio}} - \mathbb{E}[\mathcal{S}_1^{m,\text{idio}}]}{\sqrt{\text{Var}(\mathcal{S}_1^{m,\text{idio}})}} > 0\Big)$$

[24] For simplicity, we assume that interest rates are zero, or alternatively that insurance claims are already discounted.

converges to 50%, due to the central limit theorem. To stay on the safe side, a safety loading is necessary in addition to the net risk premium. Classical actuarial premium principles provide explicit safety loadings in a transparent manner, based on the first two moments of the loss distribution:

- **Expected value principle:** $\pi^{m,\text{idio}} = (1+a)\mathbb{E}[\mathcal{S}_1^{m,\text{idio}}]$ for a parameter $a > 0$,
- **Variance principle:** $\pi^{m,\text{idio}} = \mathbb{E}[\mathcal{S}_1^{m,\text{idio}}] + a\,\text{Var}(\mathcal{S}_1^{m,\text{idio}})$ with $a > 0$,
- **Standard deviation principle:** $\pi^{m,\text{idio}} = \mathbb{E}[\mathcal{S}_1^{m,\text{idio}}] + a\sqrt{\text{Var}(\mathcal{S}_1^{m,\text{idio}})}$ with $a > 0$.

The safety loading parameter a can be chosen for each cyber risk module $m = (c, k)$ separately, for example, depending on the specific loss distribution and the number of contracts n_k in tariff cell k.

In addition to these simple explicit premium principles, the safety loading can be imposed implicitly, e.g., in terms of convex principles of premium calculation including the well-known *exponential principle* or *Wang's premium principle* as special cases, cf. Example 4.1.

Systematic Risk Systematic cyber incidents affect different firms at the same time – in contrast to idiosyncratic cyber incidents. Thus, (perfect) pooling of risk is no longer applicable and classical actuarial valuation has to be replaced by a more complex analysis. In this section, we propose a valuation of systematic cyber risk in terms of modern financial mathematics, combining the principle of *risk-neutral valuation* with the theory of *monetary risk measures*, see Knispel, Stahl, and Weber (2011) for a similar discussion related to the *Market-Consistent Embedded Value* (MCEV) of insurance portfolios.

Assume that the one-year losses $\mathcal{S}_1^{m,i,\text{systematic}}$ for all policyholders i in group k with cyber risk module $m = (c, k)$ are described by frequency-severity models, and suppose that frequency and severity distributions depend on common risk factors. In this case, the total claims amount may be viewed as a contingent claim, depending on the evolution of common factors.

Generally speaking, contingent claims are contracts between two or more parties which determine future payments to be exchanged between the parties conditional or contingent on future events. Formally, a contingent claim with payoff at terminal time $t = 1$ is described as a random variable. In financial mathematics, the valuation of contingent claims relies on a financial market model on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \mathbb{P})$ with a number, say $d+1$, of liquidly traded primary products with price processes $(P_t^0)_{t\in[0,1]}, (P_t^1)_{t\in[0,1]}, \ldots, (P_t^d)_{t\in[0,1]}$. The underlying price processes can be modeled either as stochastic processes in discrete time or in continuous time. The asset '0' plays the role of a numéraire, i.e., it is used for discounting purposes. A contingent claim H_1 maturing at time $t = 1$ is called replicable or hedgeable if there exists a self-financing trading strategy[25] $\vartheta = (\vartheta_t^0, \vartheta_t^1, \ldots, \vartheta_t^d)_{t\in[0,1]}$ (specifying the number of shares ϑ_t^i of primary products in the portfolio at time t) whose terminal wealth V_1^ϑ coincides with the payoff H_1 for almost all scenarios. In absence of arbitrage, the price H_0 of a replicable contingent claim H_1 is unique and equals the *cost of perfect replication*. The calculation of this price can, however, be decoupled from the calculation of the replication strategy itself by the *principle of risk-neutral valuation*. Formally, risk-neutral valuation resembles the classical actuarial valuation in the sense that prices are computed as expectation of future payments. The real-world measure $\mathbb{P}$ must, however, be replaced by a technical probability measure $\mathbb{Q}$, called *risk-neutral measure* or *martingale measure*. The latter name is motivated by the fact that discounted prices $(P_t^i/P_t^0)_{t\in[0,1]}$, $i = 1, \ldots, d$, must be martingales with respect to $\mathbb{Q}$. The risk-neutral valuation formula yields

$$H_0 = P_0^0 \mathbb{E}_\mathbb{Q}\left[\frac{H_1}{P_1^0}\right],$$

[25] Intuitively, the self-financing condition means that the portfolio is always rearranged such that on the one hand no additional capital is required and on the other hand no capital is withdrawn.

i.e., the cost of replication can be obtained as expectation of the discounted payoff with respect to any arbitrary (equivalent) martingale measure $\mathbb{Q}$.

Markets are, however, typically incomplete[26] in the sense that not every contingent claim can be replicated in terms of liquidly traded primary products. In particular, contingent claims arising from cyber risks cannot be hedged perfectly in the financial market. For non-replicable contingent claims, risk-neutral valuation is still applicable, but now provides – depending on a whole class of martingale measures – an interval of prices which are consistent with the absence of arbitrage. Our evaluation of non-replicable contingent claims, however, is based on monetary risk measures and capital requirements.

Let us denote by $\mathcal{X}$ the set of financial positions with maturity $t = 1$ whose risk needs to be assessed. Mathematically, the family $\mathcal{X}$ is a vector space of real-valued mappings X_1 on $(\Omega, \mathcal{F}, \mathbb{P})$ that contains the constants. By sign-convention, negative values of X_1 correspond to debt or losses, i.e., the claims amount $\mathcal{S}_1^{m,i,\text{systematic}}$ corresponds to the financial position $X_1 = -\mathcal{S}_1^{m,i,\text{systematic}}$. A monetary risk measure[27] $\rho : \mathcal{X} \to \mathbb{R}$ quantifies the risk of a contingent claim that cannot be priced by the cost of perfect replication. Intuitively, a monetary risk measure can be viewed as a capital requirement: $\rho(X_1)$ is the minimal capital that has to be added to the position X_1 to make it acceptable, e.g., from the perspective of a financial supervisory authority, a rating agency, or the board of management. To capture the idea that homogeneous risks are assessed in the same way, we assume that ρ is distribution-based, i.e., $\rho(X) = \rho(Y)$ whenever the distributions of X and Y under $\mathbb{P}$ are equal. Prominent examples of distribution-based monetary risk measures are *Value at Risk* (VaR) and *Average Value at Risk* (AVaR).[28]

Combining these two approaches, an *algorithm* for the calculation of the premium $\pi^{m,\text{systematic}}$ can be summarized as follows:

1. Consider a decomposition of the financial position $-\mathcal{S}_1^{m,\text{systematic}} = H_1^m + R_1^m$, where
   - H_1^m is a replicable contingent claim with respect to the underlying market model,
   - and R_1^m denotes the residual term.
2. Calculate the premium $\pi^{m,\text{systematic}} = H_0^m + \rho(R_1^m)$, where H_0 equals the cost of perfect replication of H_1^m, and $\rho(R_1^m)$ is the cost of the risk capital for R_1^m.

The decomposition and the premium derived from it may not be unique. From the insurer's perspective, the goal of the decomposition into the summands (H_1^m, R_1^m) is the minimization of the theoretical premium $\pi^{m,\text{systematic}} = H_0^m + \rho(R_1^m)$ which provides a lower bound for the actual premium charge. The minimization problem apparently involves a trade-off between the cost of replication and the risk of the residual. In practice, it might be reasonable to impose constraints on the decomposition such as upper bounds for H_0^m and $\rho(R_1^m)$, respectively. Indeed, since the risk of the hedgeable part H_1 can be completely eliminated for the price H_0, the specification of a bound $\rho(R_1^m) \leq \rho_{\max}$ would already control the overall risk of the systematic cyber losses

[26] In absence of arbitrage, incomplete financial market models are characterized by the existence of a whole class of equivalent martingale measures.

[27] For a rigorous introduction to the theory of risk measures we refer to Föllmer and Schied (2016).

[28] For a financial position X_1, its Value at Risk at level $\lambda \in (0, 1)$ is the smallest monetary amount that needs to be added to X_1 such that the probability of a loss becomes smaller than λ:

$$\text{VaR}_\lambda(X_1) = \inf\{m \in \mathbb{R} | \mathbb{P}[X_1 + m < 0] \leq \lambda\}.$$

In particular, $\text{VaR}_\lambda(X_1) = -q^+_{X_1}(\lambda)$, where $q^+_{X_1}$ is the upper quantile function of X_1. The Average Value at Risk, also called Expected Shortfall, at level $\lambda \in (0, 1]$ is defined by

$$\text{AVaR}_\lambda(X_1) = \frac{1}{\lambda} \int_0^\lambda \text{VaR}_\alpha(X_1)\, d\alpha.$$

$\mathcal{S}_1^{m,\text{systematic}}$, in accordance with the company's risk strategy. Conversely, if the insurer's risk budget has not yet been exhausted, it might be helpful to limit the hedging expenses by a bound $H_0^m \leq H_{\max}$ and to accept the remaining risk $\rho(R_1^m)$.

This concept can be applied for each cyber risk module standalone, but provides additional benefits at portfolio level. If the underlying risk measure ρ is even subadditive (and thus provides incentives for diversification), then the lower bound for the actual premium charge can be further reduced. More precisely, for any given decomposition $-\mathcal{S}_1^{m,i,\text{systematic}} = H_1^{m,i} + R_1^{m,i}$ per cyber risk module $m = (c,k)$ and policyholder $i = 1,\dots,n_k$ in group k, the risk of the residual term of the aggregated systematic risk satisfies

$$\rho\left(\sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{i=1}^{n_k} R_1^{m,i}\right) \leq \sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\rho(R_1^{m,i}),$$

while the costs of perfect replication are additive. Thus, the total premium required at portfolio level is in fact lower than the aggregated premiums:

$$\sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{i=1}^{n_k} H_0^{m,i} + \rho\left(\sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{i=1}^{n_k} R_1^{m,i}\right) \leq \sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{i=1}^{n_k}(H_0^{m,i} + \rho(R_1^{m,i})).$$

The diversification effect can be allocated – according to the insurer's business strategy – to the cyber risk modules, yielding a reduction of the lower bound for the actual premium charge per module.

4.2. Pricing of Systemic Cyber Risks

Systemic risk is an important issue in cyber insurance. If entities are interconnected, risks may spread and amplify in cyber networks. In addition, this process depends on investments in cyber security and self protection of the agents in the network. Insurance premiums may, in turn, influence investment decisions and thereby modify the safety of the system, cf. Section 3. How to deal with complex cyber systems and the computation of systemic cyber insurance premiums is a topic of current research.

In this section, we introduce a preliminary, stylized approach that builds a bridge between the pricing of cyber insurance contracts and systemic risk measures. We consider N interconnected insurance customers in a cyber network that are also subject to idiosyncratic and systematic risk. For simplicity, we suppose that there exists only a single insurance company that offers J types of contracts. There are two dates, $t = 0$ and $t = 1$. The initial prices of the insurance contracts, represented by a matrix $M = (m_{i,j})_{i,j} \in \mathbb{R}^{N\times J}$, are $m_{i,j}$ where $i = 1,2,\dots,N$ denotes the insurance customer and $j = 1,2,\dots,J$ the contract type. Each customer i chooses a contract type j_i from this menu and is charged a premium m_{i,j_i}. Customers decide simultaneously about insurance contracts and their investments into cyber security resulting in random losses $Y^M = (Y_i^M)_{i=1,2,\dots,N}$ at date $t = 1$, the end of the considered period.

In this setting, we discuss the notion of *systemic premium principles*. Suppose that – excluding the considered cyber business – the random net asset value of the insurance firm at date $t = 1$ is given by $\tilde{E}$. Including the cyber contracts, the net asset value[29] of the insurance firm is

$$E^M \;=\; \tilde{E} \;+\; \sum_{i=1}^{N} m_{i,j_i} \;-\; \sum_{i=1}^{N} Y_i^M. \tag{7}$$

The computation of the net asset value implicitly considers network effects that influence losses and the underlying investment decisions of the insurance customers, i.e., the systemic risk inherent in the network.

[29]The interest rate over the considered period is set to 0 in this example.

Systemic premium principles[30] refer to the family of premium matrices M that are consistent with solvency requirements or risk limits and admissibility requirements of the insurance company. These can, for example, be formalized in terms of two acceptance sets[31] $\mathcal{A}^E$ and $\mathcal{A}^Y$ of monetary risk measures. The solvency condition or risk limit is satisfied, if $E^M \in \mathcal{A}^E$. An admissibility requirement is that the stand-alone business is viable, i.e.,

$$\sum_{i=1}^{N} m_{i,j_i} \; - \; \sum_{i=1}^{N} Y_i^M \; \in \; \mathcal{A}^Y. \tag{8}$$

Conditions (7) and (8) characterize the systemic premiums, i.e., the family $\mathcal{S}$ of admissible premium matrices M.

Example 4.1. *Solvency regulation varies in different regions of the world. Solvency II in the European Union and the Swiss Solvency Test in Switzerland are based*[32] *on the risk measures* VaR *and* AVaR*, respectively. These risk measures would define the acceptance set $\mathcal{A}^E$ in our setting.*

The acceptance set $\mathcal{A}^Y$, in contrast, corresponds to a classical premium principle. Important actuarial premium principles are based on convex risk measures ρ (defined w. r. t. financial positions) by choosing $\rho(-L)$ as a premium for loss position $L \in \mathcal{X} \subseteq L^0_+(\Omega, \mathcal{F})$.[33] *Suitable examples of risk measures ρ corresponding to premium principles are:*

- ***The family of entropic risk measures:***

$$\rho_\gamma(X) := \sup_{\mathbb{Q} \in \mathcal{M}_1} \{\mathbb{E}_\mathbb{Q}[-X] - \frac{1}{\gamma} H(\mathbb{Q}|\mathbb{P})\}, \quad \gamma \in (0, \infty).$$

  Here, $\mathcal{M}_1$ is the set of all probability measures on $(\Omega, \mathcal{F})$, and

$$H(\mathbb{Q}|\mathbb{P}) := \begin{cases} \mathbb{E}_\mathbb{Q}[\log \frac{d\mathbb{Q}}{d\mathbb{P}}], & \text{if } \mathbb{Q} \ll \mathbb{P}, \\ \infty, & \text{else}, \end{cases}$$

  is the relative entropy of $\mathbb{Q}$ with respect to a reference measure $\mathbb{P}$, for example, the real-world measure. Using a variational principle for the relative entropy, the entropic risk measure ρ_γ takes the explicit form $\rho_\gamma(X) = \frac{1}{\gamma} \log \mathbb{E}_\mathbb{P}[\exp(-\gamma X)]$ and thus corresponds to the exponential premium principle for the claims amount $L = -X$. Note that $\rho_\gamma(X)$ is increasing in γ and satisfies

$$\lim_{\gamma \downarrow 0} \rho_\gamma(X) = \mathbb{E}_\mathbb{P}[-X] \quad \text{and} \quad \lim_{\gamma \uparrow \infty} \rho_\gamma(X) = \operatorname{ess\,sup}(-X),$$

  i.e., the limiting cases are the negative expected value $\rho(X) = \mathbb{E}_\mathbb{P}[-X]$ (net risk premium) as a lower bound and the maximum loss as an upper bound for premium charges.

- ***Distortion risk measures:*** *For any increasing function $\psi : [0,1] \to [0,1]$ with $\psi(0) = 0$ and $\psi(1) = 1$ the map $c^\psi(A) := \psi(\mathbb{P}(A))$, $A \in \mathcal{F}$, is called a distortion of a probability measure $\mathbb{P}$. The Choquet integral*

$$\rho^\psi(X) := \int (-X) dc^\psi = \int_0^\infty c^\psi[-X > x]\, dx + \int_{-\infty}^0 (c^\psi[-X > x] - 1)\, dx$$

[30] The suggested concept of systemic premium principles parallels the notion of systemic risk measures, see Feinstein, Rudloff, and Weber (2017), Biagini et al. (2019).

[31] See Föllmer and Weber (2015) and Föllmer and Schied (2016) for reviews on monetary risk measures.

[32] To be more precise, the implementation of the regulatory rules are based on Mean-VaR and Mean-AVaR. Details are, e.g., discussed in Weber (2018), Hamm, Knispel, and Weber (2020).

[33] For details, see Section 8 in Föllmer and Knispel (2013) and the references therein.

defines a distortion risk measure, a comonotonic risk measure. If the distortion function is concave, the distortion risk measure corresponds to Wang's premium principle

$$\rho^{\psi}(X) = \int_0^{\infty} \psi(\mathbb{P}(-X > x))\, dx > \int_0^{\infty} \mathbb{P}(-X > x)\, dx = \mathbb{E}_{\mathbb{P}}[-X]$$

that guarantees a non-negative loading for any loss position $L = -X \geq 0$. In particular, the limiting case $\psi = \mathrm{id}$ again corresponds to the negative expected value which provides a lower bound for the actuarial premium.

If we introduce a weak partial order $\leq$ on $\mathbb{R}^{N,J}$ by component-wise $\leq$-comparison, the smallest admissible premiums $\bar{\mathcal{S}}$ in the family $\mathcal{S}$ of admissible premium matrices may be characterized. Although we are dealing only with one insurance firm in our specific construction, the heuristic argument of competitiveness might motivate to focus on premiums in $\bar{S}$ only. Typically, the admissible premium allocations will not be unique.

A remaining question is the choice of a specific premium allocation. Further criteria or objectives need to be specified for this purpose. We briefly discuss three options:

- **Competition:** The heuristic argument of competitiveness might also be used to argue that total premium payments should be as small as possible. This would lead to those allocations where $\sum_{i=1}^{N} m_{i,j_i}$ is minimal.
- **Competitive segments:** If some insurance customers are more price-sensitive and more important than other, one might introduce positive weights v_i, $i = 1, 2, \ldots, N$, and focus on allocations with minimal $\sum_{i=1}^{N} v_i m_{i,j_i}$.
- **Performance optimization:** If insurance customers were willing to accept any premium allocation in $\bar{\mathcal{S}}$, one could formulate an objective function of the insurance company that determines specific premium allocations. This could be a utility functional or a performance ratio such as RoRaC.

A detailed analysis of systemic premium principles in specific models and their statistical and algorithmic implementation are challenging and important questions for future research.

5. Conclusion and Future Research

In this paper, we provided a comprehensive overview of the current literature on modeling and pricing cyber insurance. For this purpose, we introduced a basic distinction between three different types of cyber risks: *idiosyncratic*, *systematic* and *systemic* cyber risks. Models for both *non-systemic* risk types were discussed within the classical actuarial framework of collective risk models. The (separate) discussion of modeling *systemic* cyber risks then focused on risk contagion among network users in interconnected environments as well as on their strategic interaction effects. Finally, we presented concepts for an appropriate pricing of cyber insurance contracts that crucially relies on the specific characteristics of each of the three risk types.

For both practitioners and academic researchers, modeling and pricing cyber insurance constitutes a relatively new topic. Due to its relevance, the area of research is growing rapidly, but modeling and pricing approaches are still in its infancy. In our analysis, we identified the following promising avenues for future research:

- **Data:** Classical actuarial approaches heavily rely on claims data. Epidemic network models, in turn, require connectivity data for the design of realistic network topologies as well as epidemic spread data for determining the values of the epidemic parameters (such as the transition rates τ and γ). Up to now, data is scarce, and in the actuarial context, often is inaccessible due to confidentiality issues. If more data becomes available,

the presented modeling approaches could more easily be tested and evaluated. Hence, building (open access) data collections for cyber risks constitutes an important issue for future research.

- **Epidemic model solutions:** Solving epidemic models becomes intractable for realistic network sizes. Therefore, modelers typically rely on Monte Carlo simulations or moment closures. However, Monte Carlo simulations are often computationally intensive, and moment closures often lack estimates of the resulting approximation error. For the design and implementation of appropriate control measures, improvement in these areas as well as the development of exact versions of the master equations are thus highly desirable.

- **Top-down approaches:** To determine the impact of global cyber epidemics on insurance portfolios, there currently exist a few top-down approaches, which solely take a population-based approach and ignore the underlying structure of the network. However, the network topology, e.g., centrality or clustering effects, possesses a significant impact on the epidemic spread. Hence, more sophisticated refinements of these models are desirable, building bridges between the bottom-up network modeling and top-down population-based approaches.

- **Dynamic strategic interaction:** Currently, the analysis of strategic interaction effects takes place in static environments, thereby neglecting systemic spread effects. Studying strategic interaction of network participants within a dynamical set-up could thus help to determine the impact of cyber insurance contracts on the policyholders' behavior and vice versa.

- **Multi-layer networks:** Manufacturing and financial transactions both heavily rely on digital technology today. Thus, cyber attacks on critical infrastructures such as supply chains or financial institutions constitute a systemic threat to modern societies. Typically, these hierarchical systems are characterized by a high degree of interdependence again. Therefore, the analysis of multi-layer networks constitutes a promising approach to modeling such systems.

- **Pricing systemic cyber risks:** In Section 4.1, we have presented a stylized approach on pricing systemic cyber risks based on the concept of systemic premium principles, a particular class of set-valued monetary risk measures. Future research should refine this idea and analyze systemic premium principles in specific models as well as their statistical and algorithmic implementation.

Finally, we would like to emphasize that this list is not exhaustive – there exists plenty of room for research on building suitable actuarial methods for modeling and pricing cyber insurance which will hopefully contribute to a more resilient and safer cyber landscape in the future.

Appendix A Classification of Cyber Risks

In this section, we present two exemplary classification approaches of cyber risks from an actuarial perspective: CRO Forum (2016) and Zeller and Scherer (2021). For general cyber classification approaches without a specific focus on insurance, we refer to the information security literature, see, e.g., Harry and Gallagher (2018) and the references therein.

CRO Forum (2016) suggests a classification by manifold factors summarized in Table 1. However, due to its granularity, it does not seem very suitable for modeling purposes. Indeed, the classification rather intents to provide a "starting point for discussion" (CRO Forum (2016), p.24) on a unifying framework for data-gathering purposes.

Table 1: CRO Forum (2016) Classification Overview

Cyber Incident	**Event Type**	**Root Causes**	**Actor**	**Impact Type**
1 System malfunctions/misuse	Operational Risk Categories	A People	1 Nation states	Business interruption
2 Data confidentiality		B External causes	2 Organised criminals	Data and software loss
3 Data integrity/availability		C Processes	3 Hackers	Theft or fraud
4 Malicious activity		D System	4 Hacktivists	Cyber ransom and extortion
			5 Insiders	Breach of privacy
				Reputational damage
				Regulatory & legal defense costs
				Fines and penalties
				Physical asset damage
				and many more, in total: 22 categories

Zeller and Scherer (2021) provides a more model-oriented classification of cyber incidents, see Table 2. The paper distinguishes between idiosyncratic and systemic incidents. However, the latter category should, in our view, be further divided into *systematic* and *systemic* incidents, see the discussion in Section 2.

Table 2: Zeller and Scherer (2021) Classification Examples (see Table 2 therein)

	Idiosyncratic incidents		*Systemic events*	
	Targeted attack	Individual failure	Untargeted attack	Mass failure
Data Breach (DB)	Targeted data theft	Individual unintended data disclosure	Data theft through widespread malware / phishing	Unintended data disclosure at cloud service provider
Business Interruption (BI)	Targeted (D)DoS / Ransomware attack	Disruption of IT system or process through accidental malfunction	Widespread ransomware attack	Cloud service outage disrupting business services
Fraud / General (FR)	CEO fraud through targeted (spear-)phishing attack	Accidental compromise of database by employee	Widespread ransomware attack or social engineering fraud	Accidental compromise of data stored at cloud service provider

Appendix B Moment Closures

This section provides details on moment closures as a measure to solve the Markovian master equation problems (5) and (6).

For node i, we let B_i be a representative of the Bernoulli random variables I_i, S_i, or R_i at a certain time t. The product $B_{j_1} \cdots B_{j_{k+1}}$, with pairwise different and ordered indices $j_1 <$

$\ldots < j_{k+1} \leq N$, is denoted by B_J, $J = \{j_1, \ldots, j_{k+1}\}$. For example, B_J with $J = \{j_1, j_2, j_3\}$ may denote a triple $I_{j_1} S_{j_2} I_{j_3}$, or $S_{j_1} S_{j_2} I_{j_3}$, etc.

A moment closure now approximates the moment $\mathbb{E}[B_J]$ by

$$\mathbb{E}[B_J] \approx H(\mathbb{E}[B_{J_1}], \ldots, \mathbb{E}[B_{J_m}]), \quad J_1, \ldots J_m \subset J, \quad |J_1|, \ldots, |J_m| \leq k.$$

Assuming that the single variables B_i are independent leads to the simplest possible moment closure, the *first order independent approximation*, also known as NIMFA in the epidemic literature[34]. It is given by

$$\mathbb{E}[B_i B_j] = \mathbb{E}[B_i]\mathbb{E}[B_j] + \mathrm{Cov}(B_i, B_j) \approx \mathbb{E}[B_i]\mathbb{E}[B_j].$$

Under this assumption, the full SIS and SIR dynamics are given by the ODE systems of equations (5) and (6), respectively, when replacing second-order moments with the corresponding product of means. The resulting systems can easily be analyzed by standard techniques from ODE theory.[35]

However, in certain network structures, the first order independent approximation may yield a large approximation error, see, e.g., Fahrenwaldt, Weber, and Weske (2018). Hence, more complex approaches to moment closures have been derived. Examples include:

1. **Split closures:** These closures are considered by Fahrenwaldt, Weber, and Weske (2018). The main idea of split closures consists in splitting a set of $k+1$ nodes into two disjoint and non-empty subsets of order $\leq k$:

   $$H(\mathbb{E}[B_{J_1}], \mathbb{E}[B_{J_2}]) = F(\mathbb{E}[B_{J_1}]) \cdot F(\mathbb{E}[B_{J_2}]), \quad J_1 \cap J_2 = \emptyset, \quad J_1 \cup J_2 = J, \quad |J_1|, |J_2| \leq k,$$

   with a *mean-field function* $F : [0,1] \to [0,1]$. Different mean-field functions lead to different approximations, e.g.:

   - **Independent approximation:** Using the identity as mean-field function, $F(y) = y$, the factorization of the moment of order $k+1$ is done as if the split components were independent:

     $$\mathbb{E}[B_J] \approx \mathbb{E}[B_{J_1}]\mathbb{E}[B_{J_2}].$$

     For the special case $k = 1$, this equals the first order independent approximation derived above.

     In the SIS model, since

     $$\mathbb{E}[I_J] = \mathbb{E}[I_{J_1}]\mathbb{E}[B_{I_2}] + \mathrm{Cov}(I_{J_1}, I_{J_2})$$

     and $\mathrm{Cov}(I_{I_1}, I_{I_2}) \geq 0$, cf. Cator and Mieghem (2014), the independent approximation leads to an *upper bound* of infection probabilities.

   - **Hilbert approximation:** The space of square-integrable random variables forms a Hilbert space with scalar product $\langle Y, Z \rangle := \mathbb{E}[Y \cdot Z]$ and corresponding norm $\|Y\| := \sqrt{\langle Y, Y \rangle} = \sqrt{\mathbb{E}[Y^2]}$. For $Y, Z \in L^2$, the scalar product defines an angle ϕ between the elements:

     $$\langle Y, Z \rangle = \|Y\| \cdot \|Z\| \cdot \cos\phi. \tag{9}$$

     Hence, taking the mean-field function $F(y) = \sqrt{y}$, and using the idempotence of Bernoulli random variables, a moment of order $k+1$ can be split into:

     $$\mathbb{E}[B_J] \approx \sqrt{\mathbb{E}[B_{J_1}]}\sqrt{\mathbb{E}[B_{J_2}]}.$$

[34] NIMFA is short for "N-intertwined mean-field approximation", see Mieghem, Omic, and Kooij (2009) for details.

[35] For the SIS model, the linear stability condition $R_0 = \frac{\tau}{\gamma} < \frac{1}{\hat{\mu}}$ for the infection-free state of the network can be obtained, where $\hat{\mu}$ denotes the spectral radius, i.e., the largest absolute eigenvalue, of the adjacency matrix A.

Due to (9), the resulting approximation error is low, if the angle ϕ is close to 0° or 180°. This observation may be used to determine an optimal split (J_1, J_2) of a given index set J.

In the SIS model, the Cauchy-Schwarz inequality implies that the first order Hilbert approximation leads to a *lower bound* of infection probabilities.

To apply these approximations, an appropriate partition scheme (J_1, J_2) for index sets J of order $k+1$ needs to be found. For the SIS model, a first optimal split procedure for both approximation types is suggested in Fahrenwaldt, Weber, and Weske (2018), Algorithm 3.13.

2. **Kirkwood closures:** These closures constitute the main approach used in the epidemic literature. The underlying idea originates from statistical physics, precisely from the so-called BBGKY hierarchy, which describes the evolution dynamics of an interacting N-particle system, originally proposed by Kirkwood (1935): Considering a set $J \subset \mathcal{V}$ of $k+1$ nodes and the corresponding moment $\mathbb{E}[B_J]$, we only take correlations into account which are stemming from infectious transmissions *over paths of length at most* $k-1$, i.e., passing through a maximum of k nodes. This idea reflects the original statistical physics approach that particle states may be assumed to be independent, if their distance exceeds a certain critical threshold.

   Now, assuming the independence of node states which are sufficiently far apart, the Kirkwood approximation estimates the moment $\mathbb{E}[B_J]$ of Bernoulli random variables with $J = \{j_1, \dots, j_{k+1}\}$ through

$$H(\mathbb{E}[B_{J_1^1}], \dots, \mathbb{E}[B_{J_{m_1}^1}], \dots, \mathbb{E}[B_{J_1^k}], \dots, \mathbb{E}[B_{J_{m_k}^k}]) = \prod_{i=1}^{k} \prod_{\ell=1}^{m_i} \mathbb{E}[B_{J_\ell^i}]^{(-1)^{k-i}},$$

   where $J_\ell^i \subset J$ denotes a subset of size i, $i \leq k$, and $\ell \in \{1, \dots, m_i\}$, i.e., m_i denotes the number of such subsets. A detailed derivation can be found, e.g, in Section V of Singer (2004). The Kirkwood approximation can be interpreted as generalization of the following scheme:

   For $k = 1$, states of any two nodes are assumed to be independent, i.e., the approximation equals the first order independent approximation described above.

   For $k = 2$, we obtain a so-called **pair-based model**. Here, the system is closed on the level of triplets and *correlations over edges* are considered. In this case, the closure reads

$$\mathbb{E}[B_{j_1} B_{j_2} B_{j_3}] = \frac{\mathbb{E}[B_{j_1} B_{j_2}] \mathbb{E}[B_{j_1} B_{j_3}] \mathbb{E}[B_{j_2} B_{j_3}]}{\mathbb{E}[B_{j_1}] \mathbb{E}[B_{j_2}] \mathbb{E}[B_{j_3}]}.$$

   Two different cases for the node triplet $\{j_1, j_2, j_3\}$ may be considered: For *closed* triplets, i.e., triplets in which edges exist between all pairs of nodes (triangles), node states are pairwise correlated, and hence, the closure cannot be reduced. In contrast, for an *open* triplet only consisting of edges (j_1, j_2) and (j_2, j_3), the states of nodes j_1 and j_3 are assumed to be independent, and therefore, the closure may be reduced to

$$\mathbb{E}[B_{j_1} B_{j_2} B_{j_3}] = \frac{\mathbb{E}[B_{j_1} B_{j_2}] \mathbb{E}[B_{j_2} B_{j_3}]}{\mathbb{E}[B_{j_2}]}.$$

   This equals the *exact result* for $\mathbb{E}[B_{j_1} B_{j_2} B_{j_3}]$ *under the independence assumption*, using Bayes' theorem.

   Thus, in the SIR Markov model, *exact closures* can be derived when considering *cut-vertices* i, i.e., nodes connecting two otherwise disconnected subgraphs G_1 and G_2 of the

network: If i is in the susceptible state of the SIR model, the infection has not yet passed through this node. Hence, the infection processes in the subgraphs G_1 and G_2, that are connected solely through i, are independent of each other, see, e.g., Kiss et al. (2015). This result in particular applies to tree graphs, where, by definition, all nodes with degree higher than one are cut-vertices and all triplets are open with $B_{j_2} = S_{j_2}$. For tree networks, the SIR pair-based model is thus exact.

References

Allianz (2021). *Allianz Risk Barometer*. Tech. rep. Allianz Global Corporate & Specialty.

Barabási, A.-L. and R. Albert (1999). "Emergence of scaling in random networks". In: *Science* 286, pp. 509–512.

Bessy-Roland, Y., A. Boumezoued, and C. Hillairet (2021). "Multivariate Hawkes process for cyber insurance". In: *Annals of Actuarial Science* 15 (1), pp. 14–39.

Böhme, R. and G. Schwartz (2010). "Modeling Cyber-Insurance: Towards a Unifying Framework". In: *WEIS*.

Biagini, F., J. Fouque, M. Frittelli, and T. Meyer-Brandis (2019). "A unified approach to systemic risk measures via acceptance sets". In: *Mathematical Finance* 29(1), pp. 329–367.

Biener, C., M. Eling, and J. H. Wirfs (2015). "Insurability of cyber risk: An empirical analysis". In: *The Geneva Papers on Risk and Insurance-Issues and Practice* 40(1), pp. 131–158.

Böhme, R. (2005). "Cyber-Insurance Revisited". In: *WEIS*.

Böhme, R. and G. Kataria (2006). "Models and measures for correlation in cyber-insurance." In: *WEIS*. Vol. 2, p. 3.

Böhme, R., S. Laube, and M. Riek (2018). "A Fundamental Approach to Cyber Risk Analysis". In: *Variance*.

Bolot, J. and M. Lelarge (2008a). "A local mean field analysis of security investments in networks". In: *NetEcon '08: Proceedings of the 3rd international workshop on Economics of networked systems*, pp. 25–30.

Bolot, J. and M. Lelarge (2008b). "Network externalities and the deployment of security features and protocols in the internet". In: *Proc. ACM Sigmetrics, Annapolis, MD*.

Bolot, J. and M. Lelarge (2009). "Economic incentives to increase security in the Internet: The case for insurance". In: *Proceedings of the 28th Conference on Computer Communications, Rio de Janeiro, Brazil*, pp. 1494–1502.

Cator, E. and P. V. Mieghem (2014). "Nodal infection in Markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated". In: *Physical Review E* 89(5).

Cooray, K. and M. M. Ananda (2005). "Modeling actuarial data with a composite lognormal-Pareto model". In: *Scandinavian Actuarial Journal* 2005(5), pp. 321–334.

CRO Forum (2016). *CRO Forum Concept Paper on a proposed categorisation methodology for cyber risk*. Tech. rep. CRO Forum.

CSIS (2020). *The Hidden Costs of Cybercrime*. Tech. rep. Center for Strategic and International Studies (CSIS) in partnership with McAfee.

Edwards, B., S. Hofmeyr, and S. Forrest (2016). "Hype and heavy tails: A closer look at data breaches". In: *Journal of Cybersecurity* 2(1), pp. 3–14.

Eling, M. (2020). "Cyber risk research in business and actuarial science". In: *European Actuarial Journal*, pp. 1–31.

Eling, M. and J. Wirfs (2019). "What are the actual costs of cyber risk events?" In: *European Journal of Operational Research* 272(3), pp. 1109–1119.

Erdős, P. and A. Rényi (1959). "On Random Graphs I". In: *Publicationes Mathematicae Debrecen* 6, pp. 290–297.

Fahrenwaldt, M. A., S. Weber, and K. Weske (2018). "Pricing of cyber insurance contracts in a network model". In: *ASTIN Bulletin: The Journal of the IAA* 48(3), pp. 1175–1218.

Feinstein, Z., B. Rudloff, and S. Weber (2017). "Measures of systemic risk". In: *SIAM Journal on Financial Mathematics* 8(1), pp. 672–708.

Föllmer, H. and T. Knispel (2013). "Convex Risk Measures: Basic Facts, Law-invariance and beyond, Asymptotics for Large Portfolios". In: *Handbook of the Fundamentals of Financial Decision Making, Part II, Eds. L.C. MacLean and W.T. Ziemba*. World Scientific.

Föllmer, H. and A. Schied (2016). *Stochastic finance: an introduction in discrete time*. 4th ed. Walter de Gruyter.

Föllmer, H. and S. Weber (2015). "The Axiomatic Approach to Risk Measurement for Capital Determination". In: *Annual Review of Financial Economics* 7, pp. 301–337.

Gillespie, D. T. (1976). "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions". In: *Journal of Computational Physics* 22(4), pp. 403–434.

Gillespie, D. T. (1977). "Exact stochastic simulation of coupled chemical reactions". In: *The Journal of Physical Chemistry* 81(25), pp. 2340–2361.

Hamm, A.-M., T. Knispel, and S. Weber (2020). "Optimal Risk Sharing in Insurance Networks". In: *European Actuarial Journal* 10(1), pp. 203–234.

Harry, C. and N. Gallagher (2018). "Classifying Cyber Events: A Proposed Taxonomy". In: *Journal of Information Warfare* 17(3), pp. 17–31.

Herath, H. and T. Herath (2011). "Copula-based actuarial model for pricing cyber-insurance policies". In: *Insurance markets and companies: analyses and actuarial computations* 2(1), pp. 7–20.

Hillairet, C. and O. Lopez (2021). "Propagation of cyber incidents in an insurance portfolio: counting processes combined with compartmental epidemiological models". In: *Scandinavian Actuarial Journal*, pp. 1–24.

Johnson, B., A. Laszka, and J. Grossklags (2014a). "How many down? toward understanding systematic risk in networks". In: *Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security*. ACM, pp. 495–500.

Johnson, B., A. Laszka, and J. Grossklags (2014b). "The Complexity of Estimating Systematic Risk in Networks". In: *Proceedings of the 27th IEEE Computer Security Foundations Symposium*. CSF.

Kermack, W. O. and A. G. McKendrick (1927). "A contribution to the mathematical theory of epidemics". In: *Proceedings of the Royal Society of London. Series A* 115, pp. 700–721.

Khalili, M. M., P. Naghizadeh, and M. Liu (2017). "Designing Cyber Insurance Policies: Mitigating Moral Hazard Through Security Pre-Screening". In: *GAMENETS 2017*, pp. 63–73.

Kirkwood, J. G. (1935). "Statistical mechanics of fluid mixtures". In: *The Journal of Chemical Physics* 3, pp. 300–313.

Kiss, I. Z., J. C. Miller, and P. L. Simon (2017). *Mathematics of Epidemics on Networks. From Exact to Approximate Models*. Vol. 46. Interdisciplinary Applied Mathematics. Springer.

Kiss, I. Z., C. G. Morris, F. Sélley, P. L. Simon, and R. R. Wilkinson (2015). "Exact deterministic representation of Markovian SIR epidemics on networks with and without loops". In: *Journal of Mathematical Biology* 70, pp. 437–464.

Knispel, T., G. Stahl, and S. Weber (2011). "From the equivalence principle to market consistent valuation". In: *Jahresbericht der Deutschen Mathematiker-Vereinigung* 113(3), pp. 139–172.

Laszka, A., E. Panaousis, and J. Grossklags (2018). "Cyber-Insurance as a Signaling Game: Self-reporting and External Security Audits". In: *Proceedings of the 9th Conference on Decision and Game Theory for Security*, pp. 508–520.

Maillart, T. and D. Sornette (2010). "Heavy-tailed distribution of cyber-risks". In: *The European Physical Journal B* 75(3), pp. 357–364.

Marotta, A., F. Martinelli, S. Nanni, A. Orlando, and A. Yautsiukhin (2017). "Cyber-insurance survey". In: *Computer Science Review*.

Martinelli, F., A. Orlando, G. Uuganbayar, and A. Yautsiukhin (2017). "Preventing the Drop in Security Investments for Non-competitive Cyber-Insurance Market". In: *International Conference on Risks and Security of Internet and Systems*, pp. 159–174.

Martinelli, F. and A. Yautsiukhin (2016). "Security by Insurance for Services". In: *Proceedings of the 1st International Workshop on Cyber Resilience Economics*.

Mazzoccoli, A. and M. Naldi (2020). "Robustness of Optimal Investment Decisions in Mixed Insurance/Investment Cyber Risk Management." In: *Risk analysis : an official publication of the Society for Risk Analysis* 40 (3), pp. 550–564.

Mieghem, P. V. and E. Cator (2012). "Epidemics in networks with nodal self-infection and the epidemic threshold". In: *Physical Review E* 86(1).

Mieghem, P. V., J. Omic, and R. Kooij (2009). "Virus Spread in Networks". In: *IEEE/ACM Transactions on Networking* 17 (1), pp. 1–14.

Mikosch, T. (2004). *Non-Life Insurance Mathematics: An Introduction With Stochastic Processes*. Non-life insurance mathematics: an introduction with stochastic processes Bd. 13. Springer.

Naghizadeh, P. and M. Liu (2014). "Voluntary Participation in Cyber-insurance Markets". In: *Proceedings of the 2014 Annual Workshop on Economics in Information Security*.

Ogut, H., N. Menon, and S. Raghunathan (2005). "Cyber insurance and IT security investment". In: *Proceedings of the 4th Workshop on the Economics of Information Security*.

Pal, R. (2012). "Cyber-Insurance in Internet Security: A Dig into the Information Asymmetry Problem". In: *CoRR* abs/1202.0884. arXiv: `1202.0884`.

Pal, R., L. Golubchik, K. Psounis, and P. Hui (2014). "Will cyber insurance improve network security? A market analysis". In: *Proceedings of the 2014 INFOCOM, IEEE*.

Pal, R., L. Golubchik, K. Psounis, and P. Hui (2019). "Security Pricing as Enabler of Cyber-Insurance: A First Look at Differentiated Pricing Markets". In: *IEEE Transactions on Dependable and Secure Computing* 16, pp. 358–372.

Pastor-Satorras, R., C. Castellano, P. Van Mieghem, and A. Vespignani (2015). "Epidemic processes in complex networks". In: *Reviews of Modern Physics*.

Schwartz, G. and S. Sastry (2014). "Cyber-insurance framework for large scale interdependent networks". In: *Proceedings of the 3rd international conference on High confidence networked systems*, pp. 145–154.

Schwartz, G., N. Shetty, and J. Walrand (2013). "Why cyber-insurance contracts fail to reflect cyber-risks". In: *51st Annual Allerton Conference on Communication, Control, and Computing, Allerton 2013, Allerton Park & Retreat Center, Monticello, IL, USA, October 2-4, 2013*. IEEE, pp. 781–787.

Shetty, N., G. Schwartz, M. Felegyhazi, and J. Walrand (2010). "Competitive cyber insurance and Internet Security". In: *Economics of Information Security and Privacy*. Springer, US, pp. 229–247.

Shetty, N., G. Schwartz, and J. Walrand (2010). "Can competitive insurers improve network security?" In: *Acquisti, A., Smith, S., Sadeghi, A.-R. (Eds.): Proceedings of the 3rd International Conference on Trust and Trustworthy Computing, in: Lecture Notes in Computer Science, vol. 6101, Springer, Berlin, Heidelberg*, pp. 308–322.

Singer, A. (2004). "Maximum entropy formulation of the Kirkwood superposition approximation". In: *Journal of Chemical Physics* 121(8), pp. 3657–66.

Sun, H., M. Xu, and P. Zhao (2020). "Modeling Malicious Hacking Data Breach Risks". In: *North American Actuarial Journal*.

Weber, S. (2018). "Solvency II, or How to Sweep the Downside Risk Under the Carpet". In: *Insurance: Mathematics and Economics* 82, pp. 191–200.

WEF, ed. (Oct. 2016). *Understanding Systemic Cyber Risk*. World Economic Forum.

Wheatley, S., T. Maillart, and D. Sornette (2016). "The extreme risk of personal data breaches and the erosion of privacy". In: *The European Physical Journal B* 89(1), pp. 1–12.

Xu, M. and L. Hua (2019). "Cybersecurity insurance: Modeling and pricing". In: *North American Actuarial Journal* 23(2), pp. 220–249.

Yang, Z. and J. Lui (2014). "Security adoption and influence of cyber-insurance markets in heterogeneous networks". In: *Performance Evaluation* 74, pp. 1–17.

Zeller, G. and M. Scherer (2021). "A comprehensive model for cyber risk based on marked point processes and its application to insurance". In: *European Actuarial Journal*.